

<https://helda.helsinki.fi>

---

## Horizontal Gene Transfer of Phytochelatin Synthases from Bacteria to Extremophilic Green Algae

Olsson, Sanna

2017-01

---

Olsson , S , Penacho , V , Puente-Sanchez , F , Diaz , S , Eduardo Gonzalez-Pastor , J & Aguilera , A 2017 , ' Horizontal Gene Transfer of Phytochelatin Synthases from Bacteria to Extremophilic Green Algae ' , Microbial Ecology , vol. 73 , no. 1 , pp. 50-60 . <https://doi.org/10.1007/s00248-016-0848-z>

---

<http://hdl.handle.net/10138/307903>

<https://doi.org/10.1007/s00248-016-0848-z>

---

acceptedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

# Horizontal gene transfer of phytochelatin synthases from bacteria to extremophilic green algae

Sanna Olsson<sup>1,2\*</sup>, Vanessa Penacho<sup>2,3</sup>, Fernando Puente-Sánchez<sup>2</sup>, Silvia Díaz<sup>4</sup>, José Eduardo Gonzalez-Pastor<sup>2</sup> and Angeles Aguilera<sup>2</sup>

<sup>1</sup>Department of Agricultural Sciences, P.O. Box 27, 00014 University of Helsinki, Finland

<sup>2</sup>Centro de Astrobiología (CSIC-INTA), Carretera de Ajalvir Km 4, Torrejón de Ardoz, 28850 Madrid, Spain

<sup>3</sup>Bioarray, S.L., Parque Científico y Empresarial de la UMH. Edificio Quorum III. Avenida de la Universidad s/n, 03202 Elche, Alicante, Spain

<sup>4</sup> Departamento de Microbiología-III, Facultad de Biología, Universidad Complutense (UCM), Madrid, Spain

\* Corresponding author: Sanna Olsson, e-mail: [sanna.olsson@helsinki.fi](mailto:sanna.olsson@helsinki.fi), tel.: +358405622456

## **ABSTRACT**

Transcriptomic sequencing together with bioinformatic analyses and an automated annotation process led us to identify novel phytochelatin synthase (PCS) genes from two extremophilic green algae (*Chlamydomonas acidophila* and *Dunaliella acidophila*). These genes are of intermediate length compared to known PCS genes from eukaryotes and PCS-like genes from prokaryotes. A detailed phylogenetic analysis gives new insight into the complicated evolutionary history of PCS genes and provides evidence for multiple horizontal gene transfer events from bacteria to eukaryotes within the gene family. A separate subgroup containing PCS-like genes within the PCS gene family is not supported since the PCS genes are monophyletic only when the PCS-like genes are included. The presence and functionality of the novel genes in the organisms was verified by genomic sequencing and qRT-PCR. Furthermore, the novel PCS gene in *Chlamydomonas acidophila* showed very strong induction by cadmium. Cloning and expression of the gene in *Escherichia coli* clearly improves its cadmium resistance. The gene in *Dunaliella* was not induced, most likely due to gene duplication.

**Keywords:** heavy metal tolerance, cadmium, acidophile, *Chlamydomonas acidophila*, *Dunaliella acidophila*

## INTRODUCTION

Contamination of soil and water by toxic heavy metals poses a major environmental hazard due to their toxicity and accumulation behavior [1]. Thus, metals are among the most studied polluting agents. Heavy metals are usually toxic because they displace endogenous metal cofactors from their cellular binding sites, undergo aberrant capping reactions with the thiol groups of proteins and thiol containing coenzymes, and promote the formation of active oxygen species [2]. Of the various detoxification pathways activated in plants under heavy metal stress, increased synthesis of sulfur containing defense compounds, namely cysteine, glutathione (GSH) and phytochelatins (PCs), is considered to be of prime importance for their tolerance and survival [3, 4]. In addition to a variety of plants, their synthesis is also increased in algae and fungi [4-5] upon metal exposure.

Phytochelatin synthase (PCS) mediate high-affinity binding and contribute to detoxification of heavy metal ions and metalloids, such as cadmium or arsenic, by promoting the vacuolysosomal sequestration of heavy metals [6]. These compounds are enzymatically synthesized from reduced glutathione (GSH) and related thiols in a  $\gamma$ -glutamyl-cysteinyltranspeptidation reaction catalyzed by phytochelatin synthase (PCS). Genes encoding phytochelatin synthases have been found in all vascular plants as well as some algae, fungi, diatoms and invertebrates [4, 7-9]. However, studies regarding PCS gene expression levels have only been examined in a few vascular plants and, in many cases, the results have been contradictory [10-12]. In addition to the eukaryotic full-length phytochelatin synthase genes, PCS-like or half-size PCS proteins were detected initially in prokaryotes (reviewed by Rea et al. [2]). These proteins were described to have a similar N-terminal but lack the more sequence-variable C-terminal, being therefore shorter (220–237 compared to 421–506 amino acid residues). Since then, PCS-like genes have also been reported in some eukaryotes (e.g. fungi: *Saccharomyces pombe*, *Phytophthora infestans*, *Dictyostelium discoideum*; heterokonta: *Tetrahymena thermophila*; metazoan: *Caenorhabditis elegans*, *Hydra magnipapillata*, *Strongylocentrotus purpuratus*) with scattered distribution but the functionality of these transcripts has largely remained untested [13-14].

Most studies concerning heavy metals and microorganism physiology are related to environments polluted by industrial and domestic wastes. Only little is known about the toxicity of these substances in environments with a natural, non-anthropogenic source of high levels of

heavy metals, such as extreme acidic ecosystems. Acidic environments tend to contain unusually high concentrations of heavy metals, because their solubility increases markedly as the pH decreases [15]. Despite these extreme environmental conditions, a large number of prokaryotic and eukaryotic organisms have been identified to live in the presence of high concentrations of heavy metals [16-17]. Acidophiles can survive in  $\leq 1000$ -fold higher heavy metals amounts than neutrophilic microorganisms [18]. These highly metal resistant organisms seem to have multiple and/or more efficient active resistance systems than neutrophiles [19]. Modern biotechnological methods like high-throughput sequencing of non-model organisms that live in extreme habitats can help to understand the metabolism that allows these organisms to live in their environment and also enable the discovery of novel chemical components.

In a recent study on genes affected by copper excess in a strain of the microalga *Chlamydomonas acidophila* isolated from Río Tinto (one of the largest extreme acidic environments described until now), transcripts with a putative function related to stress and heavy metal detoxification were identified from a *de novo* transcriptomic assembly [20]. Several of these transcripts got a best blast hit in other organisms than green algae. One of these, a transcript annotated as a glutathione gamma-glutamyl-cysteinyl-transferase (synonymous to PCS), was hypothesized to represent a novel gene absent in other green algae. A putative PCS-like transcript was later also found in another microalga, in a *Dunaliella acidophila* strain isolated from the same environment [21]. The present study aims to: 1) Characterize these novel PCS genes by additional sequencing to recover the complete genes. 2) Analyze their protein domains and transmembrane topology. 3) Study their functionality and induction under cadmium exposure. Cadmium was selected because its immense water solubility determines its wide distribution in aquatic systems and, its relatively high mobility in comparison to other metals, makes it a metal of major concern in environmental contamination [22-24]. A phylogenetic analysis was performed in order to uncover relationships between PCS genes in plants, bacteria and green algae and to interpret their evolutionary history.

## **MATERIAL AND METHODS**

### **Algae cultivation and exposure to Cadmium**

*Chlamydomonas acidophila* strain RT46 was collected from water samples taken in 2010 at the CEM station of Río Tinto (SW Spain) [17], and isolated to grow in the presence of antibiotics, vancomycin 50 µg/mL, cefotaxime 100 µg/mL and chloramphenicol 15 µg/mL (Sigma Aldrich, MO, US), on agar plates made with 0.22 µm-filtered river water. Clonal cultures were grown in 50-ml flasks in K medium [25] at pH 2 under an illumination of ca. 70 microE s<sup>-1</sup> m<sup>-2</sup> irradiance provided by day-light fluorescent tubes, 16:8 h LD cycle at 20°C.

*Dunaliella acidophila* RT5 was also collected in 2010 at RICHI sampling station, and grown in ARTII medium (developed in our laboratory, see Online Resource 1) at pH 2. When the cultivations reached stationary stage, the cultures were exposed to a Cd solution CdCl<sub>2</sub> 2 ½ H<sub>2</sub>O (Panreac Química, Barcelona, Spain). Cultures were maintained during 1h, 3h and 24h with a final concentration of 1 µM. Cells were collected in 50 ml Falcon tubes and centrifuged for 5 min at 5000 rpm. Pellets were frozen with liquid nitrogen and stored at -80 °C until RNA extraction.

### **Nucleic acid extraction**

Total RNA was extracted with TRI Reagent® Solution (Thermo Fisher Scientific, MA, US) following manufacturer's protocol. To remove possible DNA contamination the samples were treated with DNase I (RNase free) (Thermo Fisher Scientific). DNA was extracted from cultures not exposed to cadmium using the Wizard® Genomic DNA Purification Kit (Promega Corporation, WI, US) or Purelink™ Genomic DNA Mini Kit (Thermo Fisher Scientific). DNA and RNA quality were determined spectrometrically ( $A_{260}/A_{280} > 1.8$ ) using a NanoDrop ND-1000 spectrometer (NanoDrop, DE, US).

### **Primer design**

Primers for amplification of the full-length genes (CaPCS1, CaPCS2 and DaPCS1) as well as qRT-PCR, and cloning and expression of CaPCS2 were designed using the program Primer3 ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)) with default settings. Due to high variability in the PCS transcripts, species specific primers were needed. In quantitative reverse transcription PCR (qRT-PCR) 18S rRNA [26] and actin were both used as housekeeping genes to normalize gene expression. All primers used in this study together with optimal annealing temperatures are listed in Table 1.

### **PCR amplification, cloning and sequencing**

PCR amplifications (Bio-Rad Tetrad 2, Bio-Rad, CA, US) of DNA were performed in 25 µl-reactions following the recommendation for GoTaq® DNA Polymerase (Promega Corporation). PCR settings were as follows: 2 min 95°C, 35 cycles (1 min 95°C, 1 min primer specific annealing temperature, 1 min 72°C), 10 min 72°C. The number of cycles to 40 was frequently used for obtaining stronger products. The PCR products were separated on a 1% agarose gel and visualized with ethidium bromide using 50 bp DNA ladder (Thermo Fisher Scientific) for size determination. The amplification specificity and efficiency were tested and one primer pair was discarded due to unspecific amplification. The amplified products for each primer pair were cloned with the TOPO TA Cloning kit (Thermo Fisher Scientific) and sequenced on an Applied Biosystems 3730xl DNA Analyzer (Thermo Fisher Scientific) using M13pUC primers and BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific). The sequences were edited manually with PhyDE® v1.0 [27] and primer sequences eliminated before submission to GenBank (accession numbers KT348512 - KT348515).

Occasional non-specific amplification of the target region in *Dunaliella acidophila* was detected. Strong amplification of the non-specific products was detected using a gradient annealing temperature program (1 min 95 °C, 40 cycles (1 min 95 °C, 1 min 53 °C – 57 °C, 1 min 72°C, 10 min 72 °C).

### **Quantitative reverse transcription PCR (qRT-PCR)**

For qRT-PCR, protocols established by Díaz et al. [28] were followed, except that the annealing temperature was 55°C and actin (ACT1) was used in addition to 18S as endogenous control gene. ACT1 primers were modified from the *Tetrahymena* gene TTHERM\_00190950 available on the *Tetrahymena* Genome Database (TGD) [29]. The cDNA synthesis was carried out using 3 µg RNA in 20 µl-reaction, according to the protocol supplied by 1st Strand cDNA Synthesis kit (AMV, Roche, Basel, Switzerland). All qRT-PCR reactions were carried out in an iQTM5 multicolor Real-Time PCR detection System (Bio-Rad) apparatus with the following cycling conditions: (i) 5 min at 95°C to denature reverse transcriptase, (ii) 40 cycles of 95°C for 30 s, 55°C for 30 s and 72°C after 20 s. Both NTC (no template control) and RT minus control were negative. The real-time dissociation curve was used to check primer specificity and to confirm the presence of a unique PCR product. Standard curves were obtained using 10-fold serial cDNA dilutions and

determining the Ct (cycle threshold) values. The standard line parameters (amplification efficiency, slope and correlation coefficient) are reported in Online Resource 2. Analysis of relative gene expression was carried out according to the Standard-curve quantification method [30] from at least four independent experiments (each performed in duplicates). Only CaPCS2 was amplified with qRT-PCR, since CaPCS1 was not in the focus of this study and reliable results would not be possible to obtain for DaPCS2 due to assumed gene duplication.

#### **Cloning and expression of the novel PCS gene (CaPCS2) from *C. acidophilus* in *E. coli***

The CaPCS2 gene was amplified and cloned in the plasmid pdr111 using the following primers: FQUpHIN and FQLoNHE, which contain the restriction sites HindIII and NheI required for subsequent subcloning. Plasmid pdr111 was a gift from D. Rudner (Harvard Medical School) and contains the hyper-SPANK promoter, upstream of the cloning site, which is inducible by IPTG. Primer FQUpHIN contains a bacterial ribosomal binding sequence to promote the expression of the CaPCS2 gene in *E. coli*. PCR reaction was carried out using cDNA as DNA template in 20 µl reaction containing: 1x PCR buffer 1 (Expand Long Template PCR System, Roche) with 17.5 mM MgCl<sub>2</sub>, 350 µM each dNTP, 300 nM each primer, 3.75 U (0.75 µl) Expand Long Template enzyme mix (Roche) and 1 µl of cDNA. Amplification was carried out with the following cycling conditions: 2 min 94°C, 39 cycles (1 min 94°C, 1.5 min 56°C, 1.5 min 72°C), 10 min 72°C.

A unique PCR fragment of approximately 850 bp corresponding to the complete ORF of CaPCS2 gene was gel purified with the QIAquick extraction gel (Qiagen, Hilden, Germany), digested with the restriction endonucleases HindIII (Roche) and NheI (Roche) and ligated into pdr111 plasmid. Chemically competent *Escherichia coli* DH5α were transformed with pdr111 plasmid containing the cloned CaPCS2 gene. Bacterial plasmids were isolated with QIAprep Spin Miniprep kit (Qiagen). The growth medium for transformed bacteria was supplemented with 50 µg ml<sup>-1</sup> ampicillin to prevent loss of the plasmid in *E. coli*.

#### **Analysis of Cadmium resistance in *E. coli* clones**

The cadmium resistance was analyzed by the minimum inhibitory concentration (MIC) using serial dilutions 0-10 of *E. coli* clones in LB solid medium (Luria Bertani medium, Laboratorios Conda, Madrid, Spain) and with growth curves by measuring the optical density at 600 nm (OD<sub>600</sub>). Bacteria carrying the plasmid pdr111 without the gene CaPCS2 cloned were used as



negative control. The MIC were determinate pipetting 4 µl of each dilution ( $0-10^{-4}$ ) onto LB solid medium containing 200-500 µM  $\text{CdCl}_2 \cdot 2 \frac{1}{2} \text{H}_2\text{O}$  (Panreac Química). The cultures were maintained during 16-24h at 37°C. For the growth curves, OD600 was measured every 60 min by using a microplate reader (Tecan Genios, Männedorf, Switzerland) at 37°C for 50 cycles (49 h) as described in Mirete et al. [31]. At least 3 different clones in LB (non-treated) and LB supplemented with 300 or 500 µM  $\text{CdCl}_2$  (treated clones) were analyzed. The pdr111 vector contains a hyper-SPANK promoter, which makes it inducible by IPTG. For this reason, MIC determination and growth curves were carried out in the presence or in the absence of 1 mM IPTG.

### **Characterization of novel genes**

To form a complete picture of the genes both genomic and transcriptomic sequences were used. Transcripts were detected from *de novo* assemblies of *C. acidophila* [20] and *D. acidophila* [21] and in addition the corresponding sequences from another assembly for *C. acidophila* (Olsson et al. unpubl.) were retrieved. Sanger sequencing on genomic DNA was used to complement the information. A flowchart (Online Resource 3) outlines the process from culture to extraction to analyses and shows which assemblies are used in the current study.

### **Protein characterization**

The nucleotide sequences were translated to amino acids using ExPASy translation tool [32]. The novel predicted proteins were further characterized using the web-server version of HMMER [33] to identify protein domains and the Phobius web-server [34] to predict transmembrane topology and signal peptides.

### **Identification of similar PCS coding sequences and phylogenetic analyses**

The novel PCS transcripts were used as queries in BLASTx searches [35] against all protein sequences available in GenBank (nr database downloaded in April 2013) and the most similar sequences were downloaded. The number of species and different isoforms in closely related species of well represented groups like land plants were limited. Sequences were aligned with Mafft [36] using alignment option FFT-NS-I. The alignment was manually edited in PhyDE® v1.0 [27] by excluding ends of the alignments which could not be confidently aligned due to length differences and ambiguities in homology assessment, including the variable C-terminal domain.

Bayesian analyses were performed with MrBayes v3.2.1 [37], applying search strategies suggested by [37-38]. Calculations of the consensus tree and of the posterior probability of clades were performed based upon the trees sampled after the chains converged. Consensus topologies and support values from the different methodological approaches were compiled and drawn using TreeGraph2 [39].

## RESULTS AND DISCUSSION

### Transcripts coding for novel PCS genes

A full-length PCS from *Chlamydomonas acidophila* was detected and named as CaPCS1. Transcripts annotated as putative phytochelatin synthases or phytochelatin synthase-like genes in *C. acidophila* and *Dunaliella acidophila* were identified as novel genes and named CaPCS2 and DaPCS, respectively. Both of them had best BLAST hits in bacteria, but there are a few eukaryotic sequences among the most similar sequences as well (Blast results are provided in Online Resource 4). The nucleotide sequences have substantial differences between them and only the translated amino acid sequences can be aligned. The genes are illustrated in Figure 1. Additional information on the introns is represented in Online Resource 5.

### Characterization of the full-length PCS CaPCS1 from *Chlamydomonas acidophila*

The transcript that was similar to known PCS from plants and green alga from *C. acidophila*, CaPCS1 is 425 amino acids long and contains four introns. Seven different isoforms of the gene were detected. In the predicted 3'UTR (238 bp) two isoforms show a 66 bp long gap, indicating alternative polyadenylation of UTR. Alternative polyadenylation of UTR (APA-UTR) is a widespread phenomenon in eukaryotic cells [40] including *Chlamydomonas reinhardtii* [41] but has not been documented in *C. acidophila*.

There are some differences between the transcripts obtained from the different assemblies used in this study. From the *de novo* assembly by Olsson et al. [20] two different transcripts were obtained of which one (KT348511) is almost identical with the genomic sequence obtained by Sanger sequencing, while the other KT447466 has a unique 58 bp long insertion. Also the sequences from the other transcriptome assembly (Olsson et al. unpubl. corresponding sequences presented in Online resource 6) show several differences, most importantly in the

5'UTR of two of the sequences (comp17619\_c0\_seq4 and comp17619\_c0\_seq6 depicted in Fig. 2). Due to these differences it is hypothesized that two copies of the gene exist. The transcript with the longest open reading frame, which was congruent with the genomic sequence (KT348511), was used for the characterization of the gene.

#### **Characterization of CaPCS2 from *Chlamydomonas acidophila***

CaPCS2 consists of four exons and three introns, with a predicted open reading frame of 285 amino acid residues. Alternative polyadenylation of the UTR is present also in this gene; in the eight isoforms covering a 1288 bp long stretch of predicted 5'UTR regions common to all isoforms alter with regions present in only some of the isoforms (Online Resource 7).

A deletion of a nucleotide 235 bp after the beginning of the last exon was detected in the transcript assembled for this study compared to an existing one [20]. Since the deletion occurs in the coding sequence it affects the open reading frame of the gene. The genomic sequence was congruent with the transcript from our previous study (GBAH01012416), which was therefore used for the characterization of the gene for this inconsistent part (depicted in Figure 1).

#### **Characterization of DaPCS from *Dunaliella acidophila***

Two isoforms of DaPCS1 (comp2951\_c3\_seq1 and comp2951\_c3\_seq2) were predicted in the *de novo* transcriptome assembly of *Dunaliella acidophila* [21] corresponding accession numbers KT348509 -KT348510). One of them has five and the other eight introns (Figure 1).

The nucleotide sequence of contig comp2951\_c3\_seq1 is longer than comp2951\_c3\_seq2 but there is an alternative stop codon TAA (instead of CAA coding alanine), which ends the open reading frame. The open reading frame of contig comp2951\_c3\_seq2 has 390 amino acid residues including a complete phytochelatin synthase domain (pfam 05023) of 293 amino acids as well as a signal peptide and inwards orientated transmembrane domain. Contig comp2951\_c3\_seq1 has a shorter open reading frame of only 238 amino acids with a truncated phytochelatin synthase domain and neither signal peptides nor transmembrane domains are detected.

Amplification of DaPCS yielded in addition to the expected amplification product unspecific amplification products, observed by gel electrophoresis. Two different products differing about 650 bp in size were cloned and sequenced (see material and methods and Online Resource 8). The sequences are otherwise identical for the overlapping parts except for two differing nucleotides, suggesting gene duplication in this haploid organism. The sequence of the larger product is congruent with the other genomic and transcriptomic sequences and thus probably represents the prevalent gene copy.

### **Gene expression of the novel CaPCS2 gene**

Gene expression analysis of the CaPCS2 gene (*Chlamydomonas acidophila*) was carried out under cadmium treatment using qRT-PCR. The relative mRNA expression levels of target gene were normalized against the levels of actin and 18S rRNA. Statistical analysis for qRT-PCR fold induction indicates that there is a significant difference between the cadmium treatment and the control (Pair Wise Fixed Reallocation Randomisation test,  $p < 0.01$ ). The gene showed a very strong induction (1275.3 fold induction,  $\pm 218.8$  SD) at three hours, which flatten out at 24 hours (47.8 fold change,  $\pm 24.1$  SD).

As discussed above, it is likely that DaPCS is duplicated, which explains the difficulties involved in the DNA amplification of this region. For this reasons qRT-PCR amplification was not performed on DaPCS. The focus of our study is on the novel PCS genes, and therefore the function of the full-length CaPCS1 gene was not investigated in detail. Furthermore, different isoforms might be differently expressed thus introducing errors to the measurements. The results by Ramos et al. [10] showed that the regulatory mechanisms of PCS expression in plants is very complex including alternative splicing and gene duplication events.

### **Expression and cadmium resistance of CaPCS2 in *E. coli* clones**

The cadmium resistance analyzed by the minimum inhibitory concentration (MIC) as explained in material and methods clearly shows that the expression of CaPCS2 gene in *E. coli* improves their resistance to Cd (Fig. 3). The effect is observed on LB solid medium with the highest tested concentration (500  $\mu$ M) of Cd. In addition, the growth curves on liquid LB of the strain expressing CaPCS2 in the presence of Cd confirm this result. The effects of Cd in a liquid medium are different from solid medium, and the improved resistance to cadmium was better observed

in LB broth with 300  $\mu$ M added Cd (Fig. 4). In spite of the hyper-SPANK promoter in the used pdr111 plasmid, which controls the expression of the inserted gene CaPCS2 and is inducible by IPTG, addition of IPTG has no significant effect on the cell growth. This could be caused by the transcription of the gene from another constitutively promoter in the plasmid, which dims out the effect of IPTG.

### **Phylogenetic analyses of phytochelatin synthases**

The results of the phylogenetic analyses show different topology from the expected one based on known species relationships. The only large clade reflecting phylogenetic relationships of organisms is the clade that contains known PCS genes from land plants and green algae. Even within this group some species, such as lotus and *Volvox carteri*, have different variants or copies of the gene, which were not always monophyletic (Fig. 5). Indeed, many plant species are known to have at least two PCS genes [10, 13]. The PCS genes from other organisms do not form monophyletic groups congruent with species phylogenies and many of the branches remain unresolved or with low support. Some smaller clades containing bacteria, metazoa or fungi get statistical support but none of them includes all the PCS genes of the group of organisms present in the analysis. Genes from cyanobacteria have a basal position in the phylogenetic tree and do not form a clade.

The full-length PCS (CaPCS1) present in *Chlamydomonas acidophila* is located in a clade containing PCS from other green algae. No similar transcripts from *Dunaliella acidophila* or other publicly available *Dunaliella* sequences were found. This could, however, be due to limited data available for the genus *Dunaliella*. The novel PCS transcripts (CaPCS2 from *Chlamydomonas acidophila* and DaPCS1 from *Dunaliella acidophila*) cluster together but are distinguished by long branches due to very different amino acid sequences between them (Fig. 5 and Online Resource 9).

### **The scattered distribution of PCS genes in nature**

For a long time, researchers have been puzzled by the high constitutive expression of PCS genes in different plant tissues [2]. This is especially peculiar if we assume the presence of high heavy metal concentrations to be an exclusively recent phenomenon caused by human activity like mining. It might be, however, that high heavy metal concentration has been a much more

severe problem for organisms early in the earth's history and especially arsenic and cadmium comprised a major threat to incipient life forms [42-43]. Indeed, all life forms have evolved strategies to cope with these heavy metals. The best-characterized metalloid resistances are encoded by the *ars* operons of Gram-negative and Gram-positive bacteria, which are non-specialized in the way that they share common organization and genes including repressors of resistance to Zn(II), Cd(II) and Pb(II) (e.g. [44]. If life did emerge in waters rich in ionic metals, resistance to these compounds would probably be a very ancient trait, as suggested by Rosen [45]. Basal plants including mosses and other bryophyte lineages were for a long time supposed not to have PCS genes [13]. Recent results show the opposite, and the presence of constitutively expressed and functional phytochelatin synthases was demonstrated in all the bryophyte lineages by Petraglia et al. [46]. The presence of PCS genes in bryophytes were assayed by western blotting, *in vitro* activity and by measuring glutathione and phytochelatin contents but the genes have not been sequenced and were therefore not included in our phylogenetic analysis. It is thus not clear if PCSs are key innovations to extreme habitats or ancestral traits, and why they haven't been downregulated or lost through evolution. One plausible explanation would be that PCS serve other physiological functions beyond cadmium or arsenic detoxification, which has been suggested by some authors [9, 14].

According to our results PCS genes probably have a bacterial origin and were subsequently inherited to different groups, in some cases multiple times. Bacterial origin is also supported by the domain structure and size of the PCS genes, the bacterial genes being smaller. It has been noted earlier that PCS genes have a non-monophyletic distribution across kingdoms [14] but a comprehensive phylogenetic analyses including PCS genes from both eukaryotes and prokaryotes has not been made earlier. It has been suggested earlier that there might be horizontal gene transfer events involved in the PCS genes in metazoa and fungi [14] and even among vascular plants [10]. Our results confirm that the evolutionary history of these genes is indeed hard to elucidate and show that prokaryotic genes have to be included to resolve their phylogeny. The complex evolution of the PCS genes involves several gene duplications and losses or independent insertions of the full-length PCS genes in plants and green algae.

The great differences in the amino acid composition and sequence length between the novel PCS genes DaPCS and CaPCS2 point to an ancient horizontal gene transfer event, although two

independent events could also be possible. Horizontal gene transfer (HGT) from bacteria can be beneficial for eukaryotic organisms to survive in adverse conditions [47] and seems to be more common in algae than in vascular plants [48]. Recent studies have emphasized horizontally acquired key genes in adaptation to extreme and rapidly changing environments in other extremophilic algae [49-50]. The novel PCS genes characterized in this study represent a likely example of genes that enable the organism to invade ecological niches otherwise unavailable, in this case an environment toxic for most other organisms.

## **CONCLUSIONS**

Novel PCS genes from two extremophilic green algae are characterized in this study. The results of the phylogenetic analysis oppose the separation of prokaryotic PCS-like genes and suggest multiple horizontal gene transfer events from bacteria to eukaryotes within the PCS gene family. Further functional and structural analyses are needed to understand in detail the role of PCS both in heavy metal detoxification and other possible functions. The gene CaPCS2 increases cadmium tolerance in bacteria but it remains to be seen if the novel genes are transferable to vascular plants since this could be of potential interest for crop science and phytoremediation.

## **ACKNOWLEDGEMENTS**

This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO) under Grant CGL-2011-22540 to AA. We acknowledge CSC - IT Center for Science Ltd. (Finland) for the allocation of computational resources. Dr. Salvador Mirete (Centro de Astrobiología CSIC-INTA) is acknowledged for assistance with the growth curves for the analysis of cadmium resistance in *E. coli*. FP-S was supported by a JAE-pre fellowship from the Spanish Consejo Superior de Investigaciones Científicas (CSIC).

## LITERATURE

1. Jain CK (2004) Metal fractionation study on bed sediments of river Yamuna, India. *Water Res* 38:569–578
2. Rea PA, Vatamaniuk OK, Rigden DJ (2004) Weeds, Worms, and More. Papain's Long-Lost Cousin, Phytochelatase. *Plant Physiol* 136:2463–2474
3. Rausch T, Wachter A. (2005) Sulfur metabolism: a versatile platform for launching defence operations. *Trends Plant Sci* 10:503–509
4. Grill E, Winnacker E-L, Zenk MH (1987) Phytochelatins, a class of heavy-metal-binding peptides from plants, are functionally analogous to metallothioneins. *Proc Natl AcadSci USA* 84:439–443
5. Mishra S, Srivastava S, Tripathi RD, Govindarajan R, Kuriakose SV, Prasad MN (2006) Phytochelatase synthesis and response of antioxidants during cadmium stress in *Bacopa monnieri* L. *Plant Physiol Biochem* 44:25–3
6. Cobbett CS (2000) Phytochelatase biosynthesis and function in heavy-metal detoxification. *Curr Opin Plant Biol* 3:211–216
7. Clemens S, Kim EJ, Neumann D, Schroeder JI (1999) Tolerance to toxic metals by a gene family of phytochelatase synthases from plants and yeast. *EMBO J* 18:3325–3333
8. Vivares D, Arnoux P, Pignol D (2005) A papain-like enzyme at work: native and acylenzyme intermediate structures in phytochelatase synthesis. *Proc Natl Acad Sci USA* 102:18848–18853
9. Rea PA (2012) Phytochelatase: of a protease a peptide polymerase made. *Physiol. Plant* 145:154–164



10. Ramos J, Clemente MR, Naya L et al (2007) Phytochelatin Synthases of the Model Legume *Lotus japonicus*. A Small Multigene Family with differential response to Cadmium and alternatively spliced variants. *Plant Physiol.* 143:1110–1118
11. Liu Z, Gu C, Chen F, Yang D, Wu K, Chen S, Jiang J, Zhang Z (2012) Heterologous expression of a *Nelumbo nucifera* phytochelatin synthase gene enhances cadmium tolerance in *Arabidopsis thaliana*. *Appl Biochem Biotechnol* 166:722–734
12. Shukla D, Kesari R, Mishra S et al. (2012) Expression of phytochelatin synthase from aquatic macrophyte *Ceratophyllum demersum* L. enhances cadmium and arsenic accumulation in tobacco. *Plant Cell Rep* 31:1687–1699
13. Clemens S (2006) Evolution and function of phytochelatin synthases. *J Plant Physiol* 163:319–332
14. Clemens S, Peršoh D (2009) Multi-tasking phytochelatin synthases. *Plant Science* 177:266–271
15. Nordstrom DK, Alpers CN (1999) Negative pH, efflorescent mineralogy and consequences from environmental restoration at the Iron Mountain Superfund site, California. *Proc Nat Acad Sci USA* 96:3455–3462
16. Nixdorf B, Wollmann K, Deneke R (1998) Ecological potentials for planktonic development and food web interactions in extremely acidic mining lakes in Lusatia. In: Geller J, Klapper H, Salomons W (eds) *Acidic mining lakes*. Springer, New York, pp 147–167
17. Aguilera A, Manrubia SC, Gómez F, Rodríguez N, Amils R (2006) Eukaryotic community distribution and their relationship with the water physicochemical parameters in an extreme acidic environment, Río Tinto (SW, Spain). *Appl Environ Microbiol* 72:5325–5330

18. Dopson M, Baker-Austin C, Koppineedi PR, Bond P (2003) Growth in sulfidic mineral environments: metal resistance mechanisms in acidophilic micro-organisms. *Microbiol* 149:1959–1970
19. Dopson M, Ossandon FJ, Lövgren L, Holmes, D (2014) Metal resistance or tolerance? Acidophiles confront high metal loads via both abiotic and biotic mechanisms. *Front Microbiol* 5:1–4
20. Olsson S, Puente-Sánchez F, Gómez-Rodríguez M, Aguilera A (2015) Transcriptional response to copper excess and identification of genes involved in heavy metal tolerance in the extremophilic microalga *Chlamydomonas acidophila*. *Extremophiles* 19:657–672
21. Puente-Sánchez F, Olsson S, Aguilera A (2016) Comparative transcriptomic analysis of the response of *Dunaliella acidophila* (Chlorophyta) to short-term Cadmium and chronic natural metal-rich water exposures. *Microb Ecol* DOI: 10.1007/s00248-016-0824-7
22. Guanzon NG, Nakahara H, Yoshida Y (1994) Inhibitory effects of heavy metals on growth and photosynthesis of three freshwater microalgae. *Fish Sci* 60:379–384
23. Worms IA, Boltzman J, Garcia M, Slaveykova VI (2012) Cell-wall-dependent effect of carboxyl-CdSe/ZnS quantum dots on lead and copper availability to green microalgae. *Environ Pollut* 167:27–33
24. de Abreu FC, da Costa PN, Brondi AM et al (2014) Effects of cadmium and copper biosorption on *Chlorella vulgaris* Bull *Environ Contam Toxicol* 93:405–9
25. Keller MD, Selvin RC, Claus W, Guillard RRL (1987) Media for the culture of oceanic ultraphytoplankton. *J Phycol* 23:633–638
26. Fei X, Eriksson M, Li Y, Deng X (2010) A Novel Negative Fe-Deficiency-Responsive Element and a TGGCA-Type-Like FeRE Control the Expression of FTR1 in *Chlamydomonas reinhardtii*. *J Biomed Biotechnol* 2010:790247

27. Müller K, Quandt D, Müller J, Neinhuis C (2005) PhyDE ® 0.995: Phylogenetic Data Editor, <http://www.phyde.de>
28. Díaz S, Amaro F, Rico D et al (2007) *Tetrahymena metallothioneins* fall into two discrete subfamilies. PLoS ONE 2:e291
29. Stover NA, Krieger CJ, Binkley G et al (2005) Tetrahymena Genome Database (TGD): a new genomic resource for *Tetrahymena thermophila* research. Nucleic Acids Res 34:D500-D503
30. Larinov A, Krause A, Miller W (2005) Standard curve based method for relative real time PCR data processing. BMC Bioinformatics 6:62
31. Mirete S, Mora-Ruiz MR, Lamprecht-Grandío M, de Figueras CG, Rosselló-Móra R, González-Pastor JE (2015) Salt resistance genes revealed by functional metagenomics from brines and moderate-salinity rhizosphere within a hypersaline environment. Front Microbiol 6:1121.
32. Artimo P, Jonnalagedda M, Arnold K et al (2012) ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40:W597–W603
33. Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37
34. Käll L, Krogh A, Sonnhammer ELL (2007) Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. Nucl Acids Res 35:W429–32
35. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucl Acids Res 25:3389–3402
36. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucl Acids Res 30:3059–3066

37. Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574
38. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294:2310–2314
39. Stöver BC, Müller, KF (2010) TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics* 11:7
40. Di Giammartino DC, Nishida K, Manley JL (2011) Mechanisms and consequences of alternative polyadenylation. *Mol Cell* 43:853–866
41. Shen Y, Liu Y, Liu L, Liang C, Li QQ (2008) Unique features of nuclear mRNA poly(A) signals and alternative polyadenylation in *Chlamydomonas reinhardtii*. *Genetics* 179:167–176
42. Oremland RS, Saltkov CW, Wolfe-Simon F, Stolz JF (2009) Arsenic in the evolution of earth and extraterrestrial ecosystems. *Geomicrobiol J* 26:522–536
43. Dani SU (2010) Gold, coal and oil. *Med Hypotheses* 74:534–541
44. Mendoza-Cózatl DG, Jobe TO, Hauser F, Schroeder JI (2011) Long-distance transport, vacuolar sequestration, tolerance, and transcriptional responses induced by cadmium and arsenic. *Curr Opin Plant Biol* 14:554–562
45. Rosen BP (1999) Families of arsenic transporters. *Trends Microbiol* 7:207–212
46. Petraglia A, De Benedicti M, Degola F et al (2014) The capability to synthesize phytochelatins and the presence of constitutive and functional phytochelatin synthases are ancestral (plesiomorphic) characters for basal land plants. *J Exp Bot* 65:1153–1163

47. Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet* 9:605–618
48. Richards TA, Soanes DM, Foster PG, Leonard G, Thornton CR, Tallbot NJ (2009) Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *Plant Cell* 21:1897–1911
49. Blanc G, Agarkova I, Grimwood J, Kuo A, Bruggeman A, Dunigan DD et al (2012) The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biol* 13:R39
50. Schönknecht G, Chen WH, Ternes CM et al (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* 339:1207–10

## TABLES

Table 1. Oligonucleotide primers used in the study. For each region, all forward (F) and reverse (R) primers are indicated as well as optimal annealing temperature, target species and references if not designed for this study. Primers used for quantitative reverse transcription PCR are marked with an asterisk \*. Primers used for cloning and expression of the novel PCS gene in *Chlamydomonas acidophila* (CaPCS2) are marked with an exclamation mark †.

## FIGURE LEGENDS

**Fig. 1** Structures of the predicted genes DaPCS1, CaPCS1 and CaPCS2. Composition of exons (black, numbered E1, E2, etc.) including untranslated regions (gray) and introns (white) of the three genes is given in bp and drawn to scale. An asterisque (\*) indicates that only transcriptomic data is available for the UTR

**Fig. 2** A partial sequence alignment of the predicted 3'UTR in CaPCS1 from *Chlamydomonas acidophila*. The sequences from different assemblies (2227\_isotig04503 from the assembly by Olsson et al. 2015 and comp17619\_c0\_seq4 and comp17619\_c0\_seq6 from a Trinity assembly based on merged data from 454- and Illumina sequencing) show differences in both nucleotide composition and length. Clustal format uses special characters to indicate that all residues in the column are identical (\*), conserved (:) or semi-conserved substitutions have been observed (.)

**Fig. 3** A) Drop assay using serial dilutions 0-10<sup>-4</sup> of *E. coli* DH5  $\alpha$  negative control and *E. coli* DH5 $\alpha$  with CaPCS2 gene cloned into pdr111 vector in LB plate supplemented with 50  $\mu$ g ml<sup>-1</sup> ampicillin. B) *E. coli* DH5 $\alpha$  negative control in LB with 500  $\mu$ M Cd and 50  $\mu$ g ml<sup>-1</sup> ampicillin. C) *E. coli* with CaPCS2 gene cloned into pdr111 vector in LB with 500  $\mu$ M Cd and 50  $\mu$ g ml<sup>-1</sup> ampicillin.

**Fig. 4** Growth curves of *Escherichia coli* in LB broth supplemented with 300  $\mu$ M of Cd in presence (+IPTG) or absence (-IPTG) of 1mM IPTG. NC= *Escherichia coli* negative control. CaPCS2= *Escherichia coli* cells carrying plasmids with CaPCS2 gene.

**Fig.5** Phylogenetic relationships based on the novel PCS-like genes and the most similar genes obtained with a BLASTX search against NCBI's databases. The tree represents the majority consensus of trees sampled after stationarity in the Bayesian analysis. Posterior probability values from the Bayesian inference equal or greater than 0.50 are shown above each branch. The tree is unrooted. The amino acid alignment was 383 aa long. The kingdoms of each clade are indicated. A miniature phylogram of the consensus tree is depicted to show the branch lengths

#### ONLINE RESOURCE CAPTIONS

**Online Resource 1** ART II medium used for the cultivation of *Dunaliella acidophila* strain RT5. The volume was brought to 1l with distilled water and the pH adjusted with 0.1N H<sub>2</sub>SO<sub>4</sub>.

**Online Resource 2** Quantitative reverse transcription PCR standard-curve parameters for target gene (novel phytochelatin synthase in *Chlamydomonas acidophila* CaPCS2) and the expression control (housekeeping) genes 18S rRNA and actin. E= amplification efficiency, S= slope, R<sup>2</sup>= correlation coefficient.

**Online Resource 3** Flow chart on the process from culture to extraction to analyses for the sequences used in the current study.

**Online Resource 4** Results for BLASTx search using the open reading frame of transcripts coding for CaPCS2 (isotig12422) and DaPCS (comp2951\_c3\_seq1 and comp2951\_c3\_seq2) as query.

**Online Resource 5** Introns in the predicted PCS genes from *C. acidophila* and *D. acidophila*. There are three spliceosomal introns in CaPCS2, four in CaPCS1 and three in DaPCS1. The introns are either the most common GT-AG splice site or variations of it.

**Online Resource 6** Different contigs of component comp17619\_c0 coding for the CaPCS1 gene in *Chlamydomonas acidophila*.

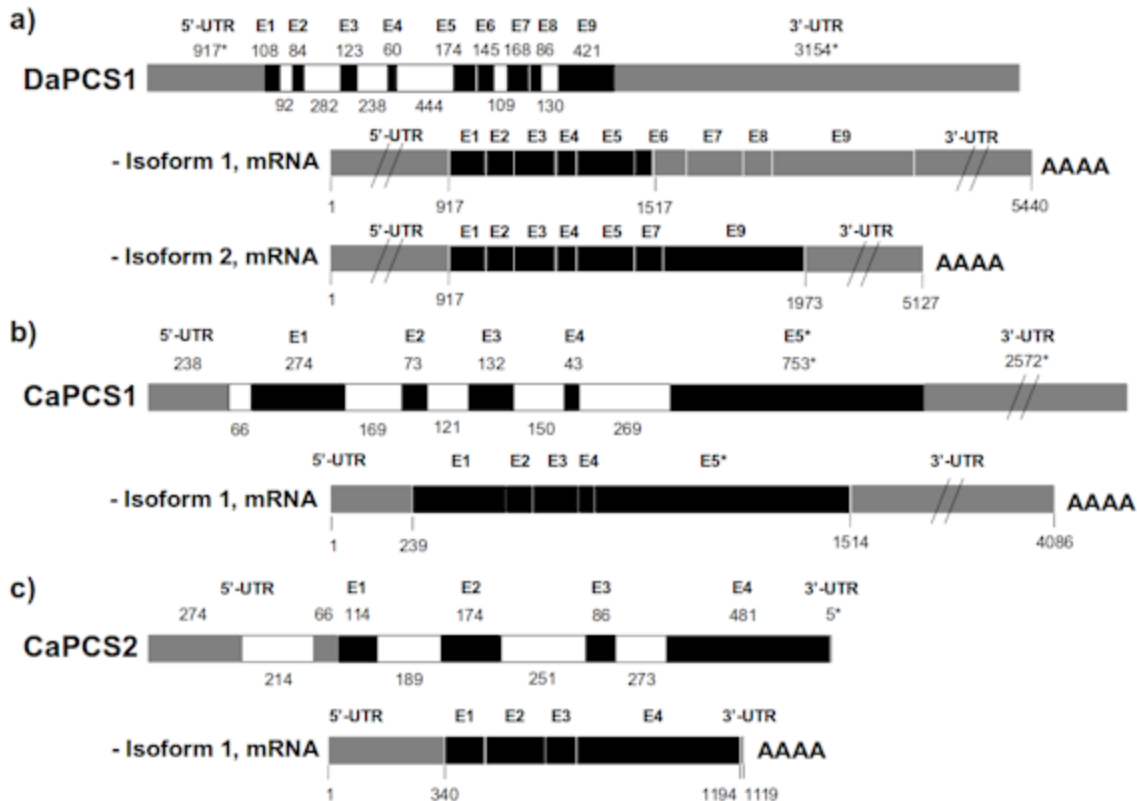
**Online Resource 7** Different contigs of component comp11852\_c0 coding for the CaPCS2 gene in *Chlamydomonas acidophila*.

**Online Resource 8** PCR products amplified in *Dunaliella acidophila* strain RT5 using gradient annealing run on 1% agarose gel (1X TBE). The gel electrophoresis shows amplification of products sized 1198 bp and 542 bp and several non-specific amplification products. The intensity of the non-specific bands depends on the used annealing temperature and PCR program. Wells 1: Ladder 1kb (Promega Corporation), marker sizes from 250 kb to 10000 kb, well 2: annealing 52-55°C. Well 3: annealing 55-60°C.

**Online Resource 9** Phylogram based on the novel phytochelatin synthase genes and the most similar genes obtained with a BLASTx search against NCBI's databases. The tree represents the majority consensus of trees sampled after stationarity in the Bayesian analysis. PP values from the Bayesian inference > 0.50 are indicated above each branch. The amino acid alignment was 383 aa long.



Primer name	5' Sequence 3'	F/R	Ann. °C	Target species	Author
CaPCH2_1F	TCACACGAAGGACAAAAGGC	F	57	<i>C. acidophila</i>	This study
CaPCH2_2F	CCAGGGTCATCACTCTCCAA	F	53	<i>C. acidophila</i>	This study
CaPCH2_1R	TGCAATTGGAGAGTGATGACC	R	57	<i>C. acidophila</i>	This study
CaPCH2_3F*	TGGGATTGGGATATTGTGCT	F	55	<i>C. acidophila</i>	This study
CaPCH2_2R*	ATCTGTTTATGCCCTGCAC	R	55	<i>C. acidophila</i>	This study
CaPCH2_3R	AGTGTCAATCAATGATGTTGCAG	R	57	<i>C. acidophila</i>	This study
CaPCH1_1F	TGAATACAACAGCGCCACAT	F	60	<i>C. acidophila</i>	This study
CaPCH1_1R	GCTCTCCACAAACATGCTCA	R	60	<i>C. acidophila</i>	This study
CaPCH1_2R	TCCTCCCCACTCTACACAC	R	60	<i>C. acidophila</i>	This study
DaPCH2_1F	TTGTGAAGATGAGGGCATGT	R	57	<i>D. acidophila</i>	This study
DaPCH2_2F	ACATCAGGCCCAACATAACG	F	57	<i>D. acidophila</i>	This study
DaPCH2_1R	GCTGCGTCTGAACATCTCAA	R	57	<i>D. acidophila</i>	This study
DaPCH2_3F	GCTTGTCCTTGAAGACCTG	F	57	<i>D. acidophila</i>	This study
DaPCH2_2R	TGCACGTTACATTCCTGCTC	R	57	<i>D. acidophila</i>	This study
18SrRNAF*	TCAACTTTCGATGGTAGGATAGTG	F	55	<i>D. acidophila</i> , <i>C. acidophila</i> , <i>C. reinhardtii</i>	Fei et al. 2010
18SrRNAR*	CCGTGTCAGGATTGGGTAATTT	R	55	<i>D. acidophila</i> , <i>C. acidophila</i> , <i>C. reinhardtii</i>	Fei et al. 2010
RT46_ACT1*	CTCACTCTCAACATTCCAGCAA	F	55	<i>C. acidophila</i> ,	Modified from TTHERM_00190950
RT46_ACT2*	GGGCCCCTCTCATCATACTC	R	55	<i>C. acidophila</i>	Modified from TTHERM_00190950
FQupHIN <sup>l</sup>	TCAGTAAAGCTTAAAGGTGGTGAAGTACTATGC AGGACCCTGCTTCA	Left	58	<i>C. acidophila</i>	This study
FQ_Lo_NHE <sup>l</sup>	ATCGTAGCTAGCAACAGCGCACAGTGGTGATA	Right	58	<i>C. acidophila</i>	This study



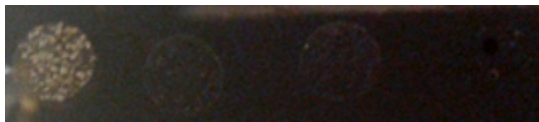
RT46\_2227\_Ctg\_i CATAAACATGGCCTTCAAGAACTTGCCAGACCAAGGATGACAATGACAATCATTTCACA  
comp17619\_c0\_se CATAAACATGGCCTTCAAGAACTTGCCAGACCAAGGATGACAATGACAATCATTTCACA  
comp17619\_c0\_se CTTAAACATGGCCTTCAAGAACCTTGCCAGACCTAG---GACAATGACACACATTTCACA  
\* \*\*\*\*\*  
  
RT46\_2227\_Ctg\_i ACGCCCTAAGCAGTCATTTACGACGCTCTAAGCAGTCGTTGTAAACTTGTAACAATAC  
comp17619\_c0\_se ACGCCCTAAGCAGTCATTTACGACGCTCTAAGCAGTCGTTGTAAACTTGTAACAATAC  
comp17619\_c0\_se ACGCCCTAAGCAGTCGTTG-----TAAACAATA-  
\*\*\*\*\*.\*. \*.\*.\*\*\*\*\* \* .\* \*\*  
  
RT46\_2227\_Ctg\_i GGCTCTTATTAGTC  
comp17619\_c0\_se GGCTCTTATTAGTC  
comp17619\_c0\_se --CTCTTATTACTC

$10^0$     $10^{-1}$     $10^{-2}$     $10^{-3}$     $10^{-4}$



*E. coli* negative control

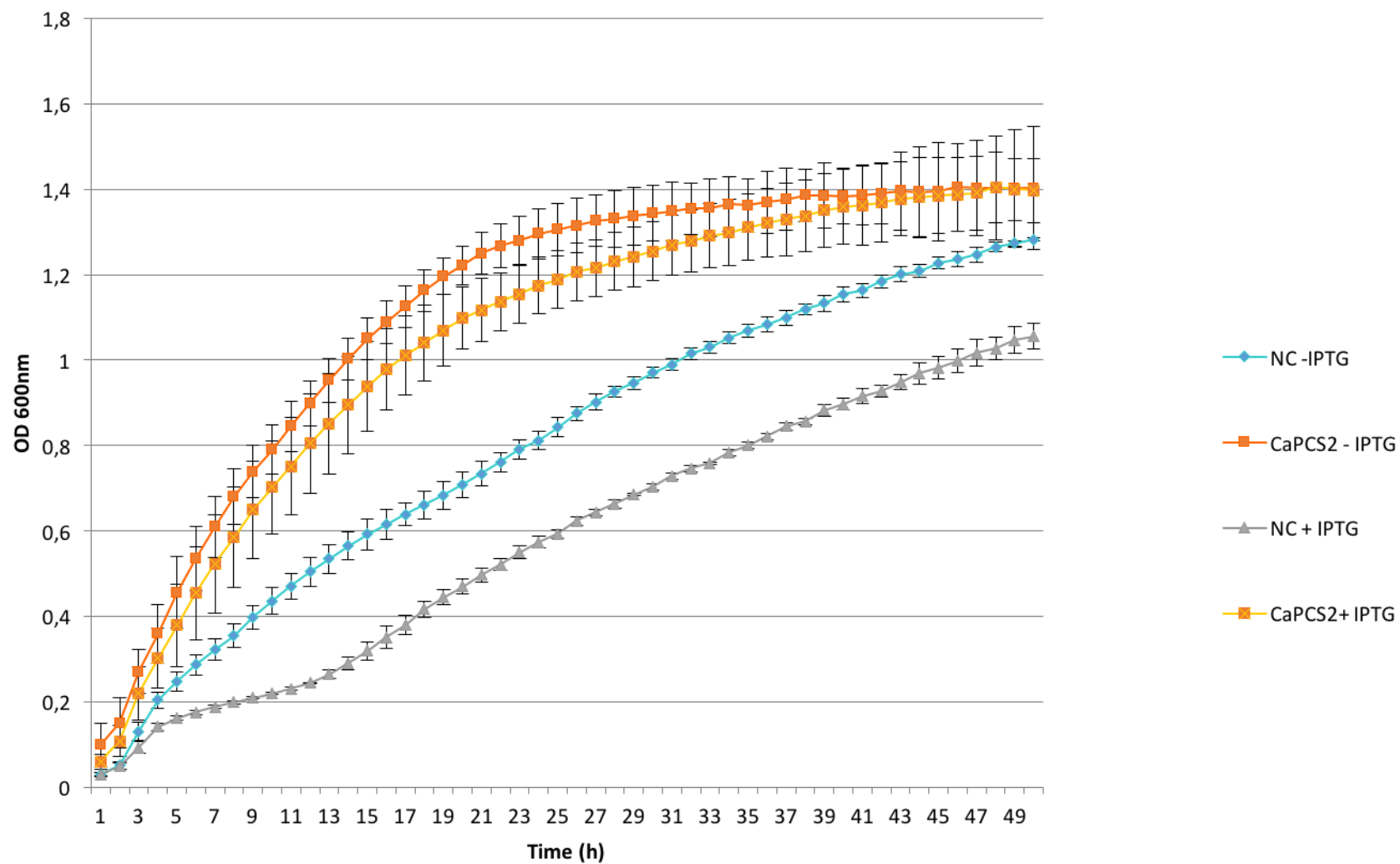
*E. coli* CaPCS2

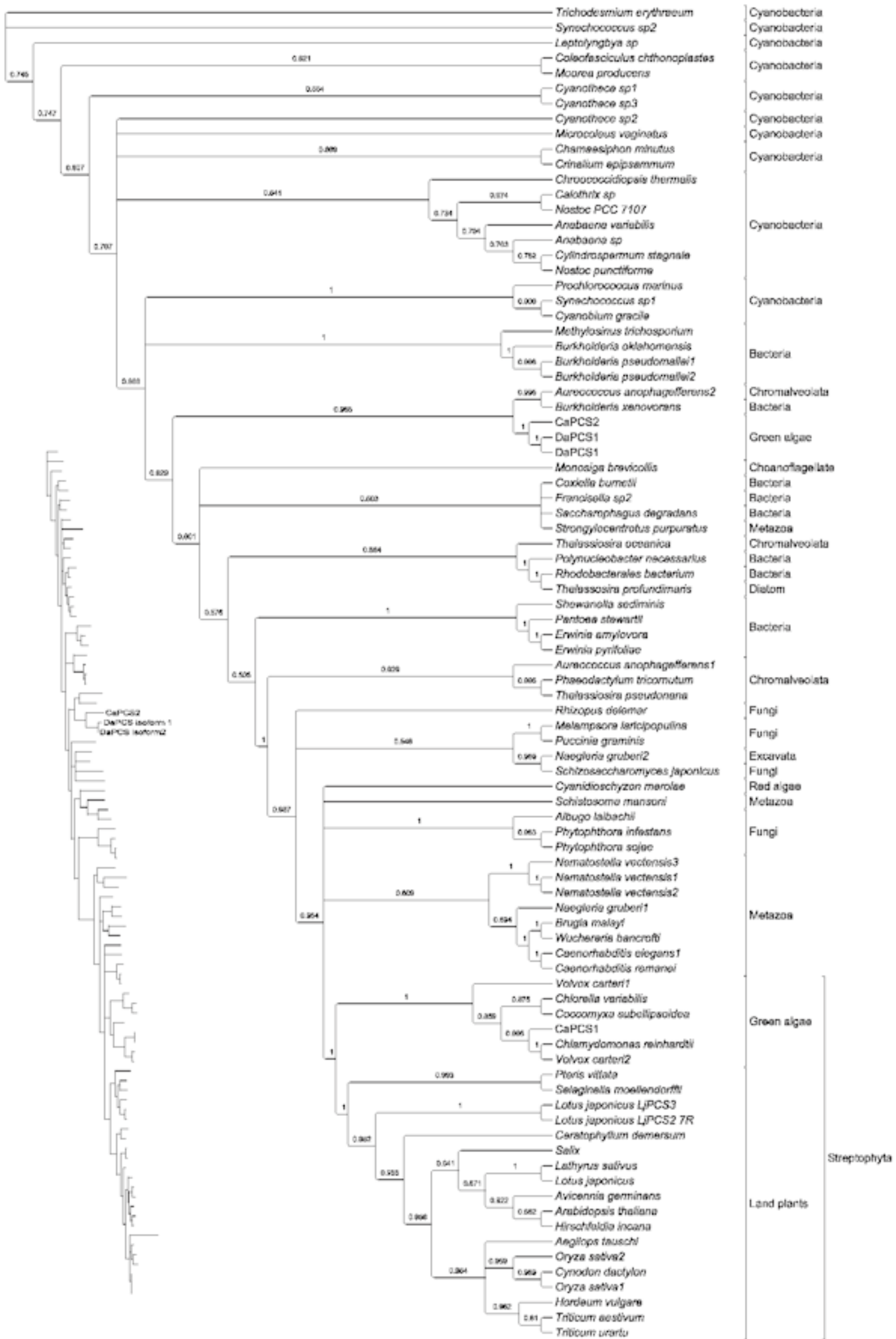


*E. coli* negative control (500  $\mu$ M Cd)



*E. coli* CaPCS2 (500  $\mu$ M Cd)





Horizontal gene transfer of phytochelatin synthases from bacteria to extremophilic green algae; Microbial Ecology; Olsson S\*, Penacho V, Puente-Sánchez F, Díaz S, Aguilera A; \* corresponding author, Department of Agricultural Sciences, P.O. Box 27, 00014 University of Helsinki, Finland, sanna.olsson@helsinki.fi

**Online Resource 1** ART II medium used for the cultivation of *Dunaliella acidophila* strain RT5. The volume was brought to 1l with distilled water and the pH adjusted with 0.1N H<sub>2</sub>SO<sub>4</sub>.

Stock Solution	ml/l
0.5M ZnSO <sub>4</sub> ·7H <sub>2</sub> O	3.5 ml
0.5M CuSO <sub>4</sub> ·5H <sub>2</sub> O	8.1 ml
0.5M Mn(NO <sub>3</sub> ) <sub>2</sub> ·6H <sub>2</sub> O	2.0 ml
0.5M Mg SO <sub>4</sub> ·7H <sub>2</sub> O	27 ml
0.5M FeSO <sub>4</sub> ·7H <sub>2</sub> O	30 ml
0.5M NiSO <sub>4</sub> ·6H <sub>2</sub> O	80 µl
0.5M KCl·7H <sub>2</sub> O	380 µl
0.5M CoCl <sub>2</sub> ·6H <sub>2</sub> O	190 µl
0.5M KBr·5H <sub>2</sub> O	30 µl

Horizontal gene transfer of phytochelatin synthases from bacteria to extremophilic green algae; Microbial Ecology; Olsson S\*, Penacho V, Puente-Sánchez F, Díaz S, Aguilera A; \* corresponding author, Department of Agricultural Sciences, P.O. Box 27, 00014 University of Helsinki, Finland, sanna.olsson@helsinki.fi

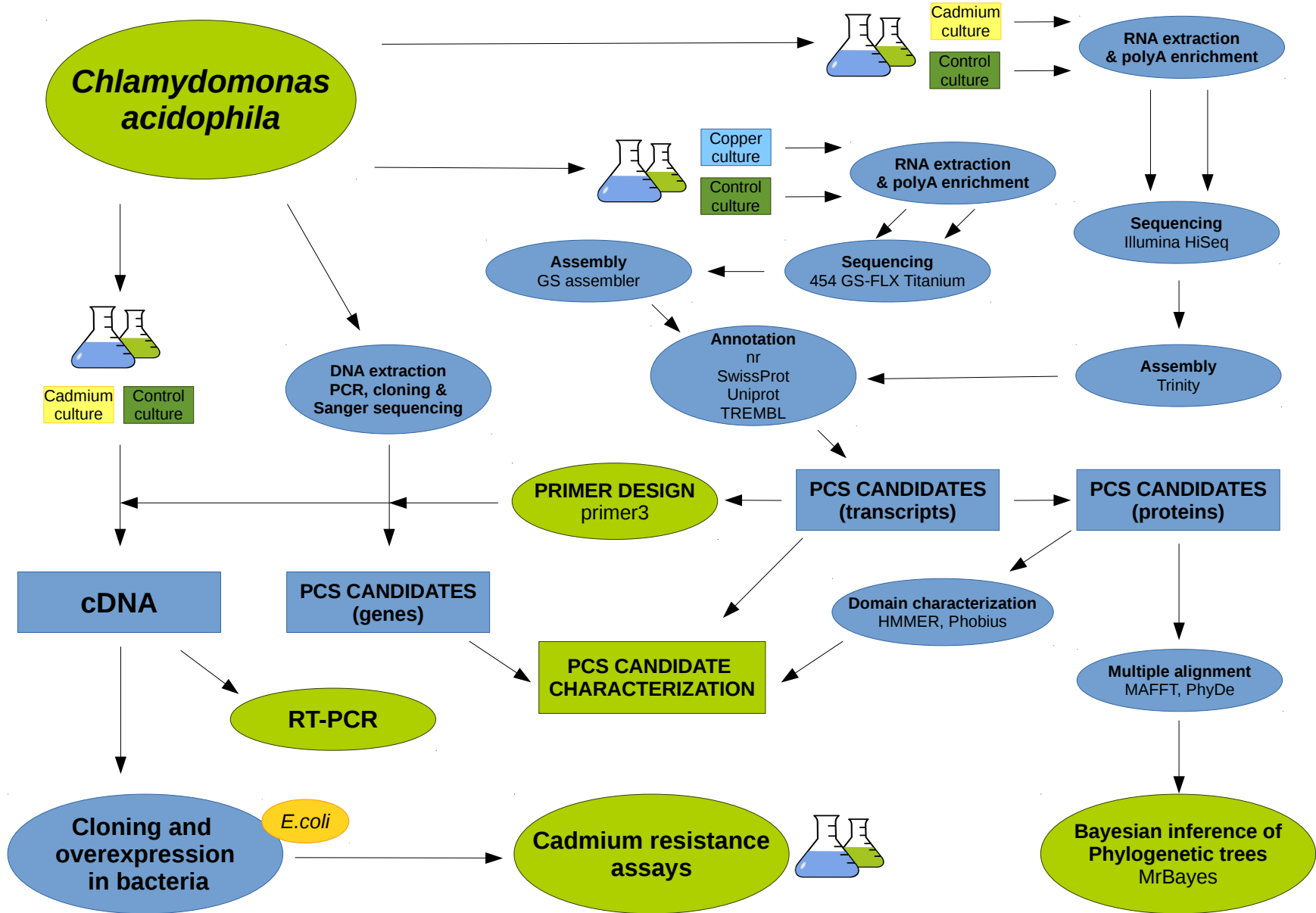
**Online Resource 2** Quantitative reverse transcription PCR standard-curve parameters for target gene (novel phytochelatin synthase in *Chlamydomonas acidophila* CaPCS2) and the expression control (housekeeping) genes 18S rRNA and actin. E= amplification efficiency, S= slope, R<sup>2</sup>= correlation coefficient.

Gene	E	S	R <sup>2</sup>
PCS	1,95	-3,43	0,99
18S rRNA	2,09	-3,12	0,98
Actin	2,10	-3,2	0,98



Horizontal gene transfer of phytochelatin synthases from bacteria to extremophilic green algae;  
Microbial Ecology; Olsson S\*, Penacho V, Puente-Sánchez F, Díaz S, Aguilera A; \* corresponding  
author, Department of Agricultural Sciences, P.O. Box 27, 00014 University of Helsinki, Finland,  
sanna.olsson@helsinki.fi

**Online Resource 3** Flow chart on the process from culture to extraction to analyses for the  
sequences used in the current study.



Electronic Supplementary Material 4. Results for BLASTx search using the open reading frame of transcripts coding for CaPCS2 (isotig12422) and DaPCS (comp2951\_c3\_seq1 and comp2951\_c3\_seq2) as query.

ORF_isotig12422	gi 821561551 ref WP_046868556.1	48.98	147	70
5 37 474 95 237 3e-27	117 ORF_isotig12422			
gi 504945810 ref WP_015132912.1	42.67	150	80	4 31 474 96
241 2e-25	112 ORF_isotig12422	gi 981461217 ref WP_059669889.1		
40.65	155 85 4 31 489 92	241 2e-24	109 ORF_isotig12422	
gi 938273386 gb KPQ08481.1	44.76	143 75 3 46 474 89 227		
3e-24	108 ORF_isotig12422	gi 981383968 ref WP_059597148.1	41.10	
146 81 3 37 474 84 224 3e-24	108 ORF_isotig12422			
gi 981251097 ref WP_059470554.1	41.10	146 81 3 37 474 93		
233 4e-24	108 ORF_isotig12422	gi 981645748 ref WP_059844321.1		
40.91	154 86 4 37 495 94	243 9e-24	107 ORF_isotig12422	
gi 705475049 gb KGS04910.1	40.52	153 84 4 37 489 83 230		
1e-23	107 ORF_isotig12422	gi 740960548 ref WP_038745121.1	40.52	
153 84 4 37 489 94 241 1e-23	107 ORF_isotig12422			
gi 782673738 ref WP_045600604.1	40.52	153 88 3 37 495 94		
243 1e-23	107 ORF_isotig12422	gi 497590677 ref WP_009904861.1		
40.52	153 88 3 37 495 94	243 1e-23	107 ORF_isotig12422	
gi 584107904 gb AHI79522.1	40.91	154 86 4 37 495 92 241		
1e-23	107 ORF_isotig12422	gi 497576287 ref WP_009890471.1	40.91	
154 86 4 37 495 94 243 1e-23	107 ORF_isotig12422			
gi 981433289 ref WP_059643234.1	40.52	153 84 4 37 489 94		
241 1e-23	107 ORF_isotig12422	gi 773064690 gb AJY42591.1	41.22	
148 84 3 31 474 82 226 2e-23	106 ORF_isotig12422			
gi 505013627 ref WP_015200729.1	38.22	157 88 3 19 471 88		
241 2e-23	106 ORF_isotig12422	gi 976471317 gb KVD74754.1	39.73	
146 83 3 37 474 84 224 2e-23	106 ORF_isotig12422			
gi 492895672 ref WP_006026078.1	39.35	155 89 3 31 495 92		
241 2e-23	106 ORF_isotig12422	gi 705585042 gb KGS47698.1	41.10	
146 83 3 37 474 84 226 3e-23	105 ORF_isotig12422			
gi 814343774 emb CPH68122.1	41.29	155 87 4 10 474 84 234		
3e-23	105 ORF_isotig12422	gi 981296166 ref WP_059513522.1	42.00	
150 84 3 25 474 89 235 4e-23	105 ORF_isotig12422			
gi 925230922 gb KOS75550.1	41.10	146 83 3 37 474 19 161		
4e-23	103 ORF_isotig12422	gi 686848709 gb KGC28119.1	41.22	148
84 3 37 480 84 228 4e-23	105 ORF_isotig12422			
gi 686840559 gb KGC20015.1	41.10	146 83 3 37 474 84 226		
4e-23	105 ORF_isotig12422	gi 976553572 gb KVE52227.1	39.73	146
83 3 37 474 84 224 4e-23	105 ORF_isotig12422			
gi 714646342 gb KGW52040.1	41.10	146 83 3 37 474 84 226		
4e-23	105 ORF_isotig12422	gi 741015566 ref WP_038797782.1	41.10	
146 83 3 37 474 94 236 4e-23	105 ORF_isotig12422			
gi 714526348 gb KGV56958.1	41.10	146 83 3 37 474 92 234		
4e-23	105 ORF_isotig12422	gi 714575209 gb KGW05595.1	41.10	146
83 3 37 474 95 237 5e-23	105 ORF_isotig12422			
gi 648412694 ref WP_026104445.1	43.62	149 78 5 34 474 99		
243 5e-23	105 ORF_isotig12422	gi 497606907 ref WP_009921091.1		
41.10	146 83 3 37 474 94	236 5e-23	105 ORF_isotig12422	
gi 740942487 ref WP_038727502.1	41.10	146 83 3 37 474 97		
239 5e-23	105 ORF_isotig12422	gi 740984489 ref WP_038768164.1		
41.78	146 82 3 37 474 94	236 5e-23	105 ORF_isotig12422	
gi 740980945 ref WP_038764816.1	41.22	148 84 3 37 480 97		
241 6e-23	105 ORF_isotig12422	gi 752526330 ref WP_041197894.1		
41.10	146 83 3 37 474 94	236 6e-23	105 ORF_isotig12422	

gi 686936961 gb KGD15055.1	41.10	146	83	3	37	474	52	194
6e-23	104	ORF_isotigl2422	gi 740959588 ref WP_038744210.1	41.10				
146	83	3	37	474	94	236	6e-23	105
gi 714507186 gb KGV37859.1	41.10	146	83	3	37	474	84	226
6e-23	105	ORF_isotigl2422	gi 490304232 ref WP_004199532.1	41.10				
146	83	3	37	474	84	226	7e-23	104
gi 484585464 gb AGK47089.1	40.54	148	85	3	31	474	82	226
7e-23	104	ORF_isotigl2422	gi 814476642 emb CPI32267.1	41.10				146
83	3	37	474	92	234	8e-23	104	ORF_isotigl2422
gi 752524486 ref WP_041196068.1	41.10	146	83	3	37	474	94	
236	9e-23	104	ORF_isotigl2422	gi 714591842 gb KGW22095.1	41.10			
146	83	3	37	474	84	226	9e-23	104
gi 754360430 ref WP_041861630.1	40.54	148	85	3	31	474	92	
236	9e-23	104	ORF_isotigl2422	gi 490685647 ref WP_004550267.1				
41.10	146	83	3	37	474	92	234	1e-22
gi 913356948 ref WP_050376444.1	41.10	146	83	3	37	474	94	
236	1e-22	104	ORF_isotigl2422	gi 800785064 emb CFL25926.1	41.10			
146	83	3	37	474	92	234	1e-22	104
gi 953804284 ref WP_058039816.1	41.10	146	83	3	37	474	94	
236	1e-22	104	ORF_isotigl2422	gi 490301683 ref WP_004197081.1				
41.10	146	83	3	37	474	94	236	1e-22
gi 953803876 ref WP_058039691.1	41.10	146	83	3	37	474	94	
236	1e-22	104	ORF_isotigl2422	gi 981741072 ref WP_059934588.1				
38.71	155	90	3	31	495	92	241	1e-22
gi 981354997 ref WP_059570263.1	39.61	154	88	3	37	498	94	
242	1e-22	104	ORF_isotigl2422	gi 751301901 ref WP_041010414.1				
38.61	158	91	4	28	492	86	240	1e-22
gi 740967447 ref WP_038751917.1	41.10	146	83	3	37	474	94	
236	1e-22	104	ORF_isotigl2422	gi 751277775 ref WP_040986624.1				
39.47	152	86	4	28	474	86	234	2e-22
gi 751301199 ref WP_041009719.1	38.61	158	91	4	28	492	86	
240	2e-22	103	ORF_isotigl2422	gi 913438795 ref WP_050418395.1				
40.26	154	87	4	13	474	88	236	2e-22
gi 909830227 gb AKU13025.1	40.26	154	87	4	13	474	93	241
2e-22	103	ORF_isotigl2422	gi 806494169 gb KKC16091.1	40.41				146
84	3	37	474	84	226	2e-22	103	ORF_isotigl2422
gi 568795567 gb AHE26011.1	41.10	146	83	3	37	474	84	226
3e-22	103	ORF_isotigl2422	gi 772972989 gb AJX33047.1	39.73				146
83	3	37	474	92	232	3e-22	103	ORF_isotigl2422
gi 756829253 ref WP_042646352.1	39.47	152	86	4	28	474	86	
234	3e-22	103	ORF_isotigl2422	gi 741019975 ref WP_038801977.1				
39.73	146	83	3	37	474	93	233	3e-22
gi 740943421 ref WP_038728427.1	40.41	146	84	3	37	474	94	
236	4e-22	103	ORF_isotigl2422	gi 740986897 ref WP_038770440.1				
41.10	146	83	3	37	474	94	236	4e-22
gi 785594494 ref WP_045719475.1	40.41	146	84	3	37	474	94	
236	4e-22	102	ORF_isotigl2422	gi 741007236 ref WP_038789710.1				
40.41	146	84	3	37	474	94	236	4e-22
gi 493029431 ref WP_006099075.1	42.36	144	77	5	46	474	104	
242	4e-22	102	ORF_isotigl2422	gi 237502694 gb ACQ95012.1	41.10			
146	83	3	37	474	198	340	5e-22	103
gi 714902087 gb KGX30928.1	41.10	146	83	3	37	474	206	348
5e-22	103	ORF_isotigl2422	gi 157806771 gb EDO83941.1	41.10				146
83	3	37	474	212	354	6e-22	103	ORF_isotigl2422
gi 160698859 gb EDP88829.1	41.10	146	83	3	37	474	190	332
7e-22	103	ORF_isotigl2422	gi 689678301 emb CEF48871.1	38.36				146
87	3	37	474	94	236	7e-22	102	ORF_isotigl2422

gi 148027253 gb EDK85274.1	41.10	146	83	3	37	474	194	336
7e-22	103	ORF_isotig12422	gi 243063500 gb EES45686.1	41.10		146		
83	3	37	474	194	336	7e-22	103	ORF_isotig12422
gi 121227363 gb ABM49881.1	41.10	146	83	3	37	474	196	338
7e-22	103	ORF_isotig12422	gi 126229014 gb ABN92554.1	41.10		146		
83	3	37	474	199	341	7e-22	103	ORF_isotig12422
gi 685679385 gb AIO66505.1	39.73	146	83	3	37	474	92	232
8e-22	102	ORF_isotig12422	gi 126219717 gb ABN83223.1	41.10		146		
83	3	37	474	203	345	8e-22	103	ORF_isotig12422
gi 169654349 gb EDS87042.1	41.10	146	83	3	37	474	209	351
8e-22	103	ORF_isotig12422	gi 184209881 gb EDU06924.1	41.10		146		
83	3	37	474	205	347	8e-22	103	ORF_isotig12422
gi 741018909 ref WP_038800916.1	39.73	146	83	3	37	474	93	
233	8e-22	102	ORF_isotig12422	gi 157936951 gb EDO92621.1	41.10			
146	83	3	37	474	214	356	8e-22	103
gi 254218068 gb EET07452.1	41.10	146	83	3	37	474	216	358
8e-22	103	ORF_isotig12422	gi 217395731 gb EEC35749.1	41.10		146		
83	3	37	474	222	364	8e-22	103	ORF_isotig12422
gi 225927831 gb EEH23872.1	41.10	146	83	3	37	474	225	367
8e-22	103	ORF_isotig12422	gi 134251017 gb EBA51096.1	41.10		146		
83	3	37	474	220	362	8e-22	103	ORF_isotig12422
gi 501733845 ref WP_012629148.1	41.18	153	79	5	31	474	95	
241	8e-22	102	ORF_isotig12422	gi 927917924 emb CDX53272.1	39.47			
152	86	4	28	474	86	234	1e-21	101
gi 499931312 ref WP_011612046.1	42.36	144	79	3	43	474	102	
241	2e-21	100	ORF_isotig12422	gi 503088795 ref WP_013323642.1				
41.98	162	81	6	13	474	87	243	3e-21
gi 826018244 ref WP_047165229.1	41.67	144	80	3	43	474	90	
229	4e-21	100	ORF_isotig12422	gi 504949901 ref WP_015137003.1				
44.52	155	76	7	25	474	93	242	5e-21
gi 685089209 emb CDX24205.1	36.71	158	94	4	28	492	86	240
7e-21	99.4	ORF_isotig12422	gi 751264288 ref WP_040974001.1					
38.16	152	88	4	28	474	88	236	8e-21
ORF_isotig12422	gi 966518160 ref WP_058534654.1	39.51	162	85				
5	13	480	75	229	9e-21	99.0	ORF_isotig12422	
gi 974600092 ref WP_059190108.1	36.71	158	94	4	28	492	86	
240	1e-20	98.6	ORF_isotig12422	gi 764950513 ref WP_044520790.1				
38.18	165	90	6	19	492	80	239	2e-20
ORF_isotig12422	gi 17130321 dbj BAB72932.1	38.18	165	90	6			
19	492	83	242	2e-20	98.2	ORF_isotig12422		
gi 652914138 ref WP_027168079.1	38.82	152	87	4	28	474	86	
234	2e-20	97.8	ORF_isotig12422	gi 701274860 gb AIV79086.1				
39.61	154	82	5	37	474	84	234	3e-20
ORF_isotig12422	gi 75702893 gb ABA22569.1	38.36	159	87	5			
19	474	80	234	3e-20	97.4	ORF_isotig12422		
gi 759580794 ref WP_043299684.1	39.61	154	82	5	37	474	94	
244	4e-20	97.4	ORF_isotig12422	gi 657935844 ref WP_029636980.1				
34.81	158	94	3	19	474	88	242	4e-20
ORF_isotig12422	gi 496440563 ref WP_009149408.1	39.33	150	85				
4	31	474	109	254	5e-20	97.4	ORF_isotig12422	
gi 504970487 ref WP_015157589.1	40.49	163	82	6	13	474	80	
236	6e-20	96.7	ORF_isotig12422	gi 505018197 ref WP_015205299.1				
37.04	162	89	4	13	474	86	242	1e-19
ORF_isotig12422	gi 748165414 ref WP_039738990.1	37.11	159	89				
5	19	474	88	242	2e-19	95.5	ORF_isotig12422	
gi 953510847 dbj BAT53535.1	35.44	158	93	3	19	474	80	234
2e-19	95.5	ORF_isotig12422	gi 501377640 ref WP_012409206.1					

32.30	161	98	3	16	474	82	239	2e-19	95.5				
ORF_isotig12422				gi 85544315 pdb 2BTW B			38.18		165	90	6	19	
492	94	253	2e-19	95.5	ORF_isotig12422		gi 85544314 pdb 2BTW A						
38.18		165	90	6	19	492	94	253	2e-19	95.1			
ORF_isotig12422				gi 737172391 ref WP_035158595.1			41.26			143	79		
4	49	474	101	239	2e-19	95.1	ORF_isotig12422						
gi 515888082 ref WP_017318665.1				35.19		162	92	5	16	474	85		
242	3e-19		94.7	ORF_isotig12422		gi 505020123 ref WP_015207225.1							
35.86		145	90	2	37	471	100	241	3e-19	94.7			
ORF_isotig12422				gi 489705594 ref WP_003609731.1			39.24			158	90		
4	1	474	86	237	3e-19	94.7	ORF_isotig12422						
gi 504891436 ref WP_015078538.1				38.10		147	86	4	37	474	100		
242	4e-19		94.4	ORF_isotig12422		gi 763312982 ref WP_044171432.1							
38.41		151	86	4	31	474	87	233	9e-19	93.2			
ORF_isotig12422				gi 658419581 ref WP_029646255.1			37.24			145	87		
3	34	468	100	240	1e-18	93.2	ORF_isotig12422						
gi 516956783 ref WP_018183684.1				38.82		152	87	4	28	474	87		
235	1e-18		93.2	ORF_isotig12422		gi 497231444 ref WP_009545706.1							
42.76		145	76	6	46	474	99	238	2e-18	92.8			
ORF_isotig12422				gi 746961585 gb KIF19114.1			42.21			154	77	8	
31	474	92	239	2e-18	92.4	ORF_isotig12422							
gi 917776332 ref WP_052290273.1				42.21		154	77	8	31	474	83		
230	2e-18		92.4	ORF_isotig12422		gi 500232799 ref WP_011902392.1							
37.41		147	83	4	31	462	87	227	3e-18	91.7			
ORF_isotig12422				gi 515356232 ref WP_016865554.1			35.29			153	93		
3	25	474	92	241	3e-18	91.7	ORF_isotig12422						
gi 751564302 ref WP_041033204.1				36.42		162	93	6	1	474	87		
242	4e-18		91.7	ORF_isotig12422		gi 656037395 ref WP_029076335.1							
35.26		156	96	3	25	492	89	239	5e-18	90.9			
ORF_isotig12422				gi 522080484 ref WP_020591693.1			38.46			143	84		
3	46	474	95	233	5e-18	90.9	ORF_isotig12422						
gi 495550750 ref WP_008275329.1				42.57		148	78	6	46	483	99		
241	6e-18		90.9	ORF_isotig12422		gi 496167609 ref WP_008892116.1							
39.16		143	83	3	46	474	91	229	6e-18	90.9			
ORF_isotig12422				gi 427346779 gb AFY29492.1			37.84			148	87	4	
34	474	81	224	6e-18	90.9	ORF_isotig12422							
gi 504925739 ref WP_015112841.1				37.66		154	88	5	25	474	93		
242	7e-18		90.9	ORF_isotig12422		gi 817130062 ref WP_046502581.1							
37.76		143	85	3	46	474	99	237	7e-18	90.9			
ORF_isotig12422				gi 1000089089 gb KXJ56520.1			39.16			143	83	3	
46	474	91	229	7e-18	90.5	ORF_isotig12422							
gi 522054471 ref WP_020565680.1				39.04		146	85	3	37	474	108		
249	7e-18		90.9	ORF_isotig12422		gi 504967338 ref WP_015154440.1							
38.16		152	86	5	31	474	94	241	1e-17	90.5			
ORF_isotig12422				gi 917319449 ref WP_051926161.1			38.19			144	84		
4	46	474	103	242	1e-17	90.5	ORF_isotig12422						
gi 746710169 ref WP_039669460.1				34.23		149	97	1	37	483	19		
166	1e-17		88.6	ORF_isotig12422		gi 759607566 ref WP_043325766.1							
37.84		148	87	4	34	474	92	235	1e-17	90.1			
ORF_isotig12422				gi 501565810 ref WP_012570258.1			34.23			149	97		
1	37	483	19	166	1e-17	88.2	ORF_isotig12422						
gi 648291774 ref WP_026072923.1				38.41		151	87	5	34	483	98		
243	2e-17		89.7	ORF_isotig12422		gi 827519218 gb KLN61905.1							
37.06		143	86	3	46	474	83	221	2e-17	89.4			
ORF_isotig12422				gi 920578875 ref WP_053005650.1			37.06			143	86		
3	46	474	99	237	2e-17	89.4	ORF_isotig12422						
gi 740234212 ref WP_038075400.1				37.16		148	88	4	37	474	92		

[illegible]

gi 754969497 ref WP_042325399.1	36.71	158	95	4	1	474	119
271 7e-15	82.4	ORF_isotig12422	gi 156379434 ref XP_001631462.1				
45.83	72	39	0	241	456	147	218 7e-15 81.6
ORF_isotig12422	gi 752594371 ref WP_041263758.1	34.69	98	61			
2 181 474 6	100	9e-15	78.6	ORF_isotig12422			
gi 738534441 ref WP_036482850.1	31.79	151	96	4	28	474	103
248 1e-14	81.3	ORF_isotig12422	gi 499442921 ref WP_011130385.1				
36.57	134	82	3	34	432	105	236 2e-14 80.5
ORF_isotig12422	gi 739030777 ref WP_036911908.1	34.81	135	83			
4 34 432 105	236	4e-14	79.7	ORF_isotig12422			
gi 499786267 ref WP_011467001.1	42.57	101	53	2	196	498	222
317 5e-14	80.5	ORF_isotig12422	gi 123963072 gb ABM77828.1				
33.78	148	93	4	34	474	105	248 6e-14 79.3
ORF_isotig12422	gi 331218212 ref XP_003321784.1	41.98	81	47			
0 196 438 150	230	7e-14	81.3	ORF_isotig12422			
gi 550281002 ref WP_022607401.1	35.62	146	84	5	49	483	99
235 9e-14	78.6	ORF_isotig12422	gi 800135199 dbj GAO34503.1				
36.55	145	88	3	40	474	90	230 2e-13 77.4
ORF_isotig12422	gi 331218210 ref XP_003321783.1	42.31	78	45			
0 202 435 277	354	2e-13	79.3	ORF_isotig12422			
gi 336447064 gb AEI36370.1	41.89	74	41	1	253	474	6 77
3e-13	73.9	ORF_isotig12422	gi 909606181 gb KNE96553.1				38.20
89 55 0	202	468	220	308	3e-13	79.7	ORF_isotig12422
gi 831778908 ref XP_012755554.1	37.84	111	64	3	181	513	166
271 3e-13	79.3	ORF_isotig12422	gi 290993528 ref XP_002679385.1				
38.82	85	52	0	202	456	145	229 4e-13 76.6
ORF_isotig12422	gi 675852512 ref XP_009011816.1	38.74	111	66			
1 211 543 206	314	6e-13	78.6	ORF_isotig12422			
gi 676386555 ref XP_009036556.1	41.76	91	52	1	205	474	152
242 7e-13	76.3	ORF_isotig12422	gi 504984674 ref WP_015171776.1				
37.90	124	75	2	145	513	111	233 9e-13 77.8
ORF_isotig12422	gi 568027492 gb ETM01446.1	37.00	100	60	2		
175 474 59	155	1e-12	75.1	ORF_isotig12422			
gi 999977806 gb KXJ16980.1	41.77	79	44	1	238	474	430 506
1e-12	77.8	ORF_isotig12422	gi 999977806 gb KXJ16980.1				32.61
138 67 5	142	492	138	270	5e-12	75.9	ORF_isotig12422
gi 909134564 gb KNE60758.1	33.61	119	76	2	145	501	24 139
1e-12	75.5	ORF_isotig12422	gi 586633751 gb EWS64201.1				39.18
97 55 1	199	489	31	123	1e-12	73.2	ORF_isotig12422
gi 999977804 gb KXJ16978.1	34.65	101	58	2	190	474	152 250
1e-12	75.9	ORF_isotig12422	gi 737401579 ref WP_035382755.1				
35.33	150	89	4	37	480	86	229 2e-12 75.1
ORF_isotig12422	gi 685951299 ref XP_009269521.1	38.54	96	57			
2 145 432 127	220	2e-12	77.0	ORF_isotig12422			
gi 524911978 ref XP_005110845.1	35.87	92	59	0	145	420	116
207 2e-12	77.0	ORF_isotig12422	gi 198417053 ref XP_002128372.1				
33.07	127	84	1	145	525	110	235 2e-12 76.6
ORF_isotig12422	gi 219124481 ref XP_002182531.1	30.63	160	83			
3 16 456 393	537	3e-12	76.6	ORF_isotig12422			
gi 909134565 gb KNE60759.1	33.61	119	76	2	145	501	93 208
3e-12	75.1	ORF_isotig12422	gi 909134566 gb KNE60760.1				33.61
119 76 2	145	501	119	234	4e-12	75.5	ORF_isotig12422
gi 727439853 ref XP_010501307.1	37.63	93	58	0	142	420	108
200 4e-12	76.6	ORF_isotig12422	gi 727439853 ref XP_010501307.1				
37.63	93	58	0	142	420	313	405 4e-12 76.6
ORF_isotig12422	gi 909138881 gb KNE64222.1	33.61	119	76	2		
145 501 119	234	5e-12	75.9	ORF_isotig12422			



gi 564587188 gb AHB86971.1	41.56	77	45	0	190	420	124	200
6e-12	75.9	ORF_isotig12422	gi 223995009 ref XP_002287188.1					
51.67	60	29	0	253	432	144	203	7e-12
72.8								
ORF_isotig12422	gi 505014688 ref WP_015201790.1	38.83		103	58			
3	169	474	118	216	7e-12	75.1	ORF_isotig12422	
gi 953489559 emb CEG46925.1	36.54	104	63	2	163	474	265	365
7e-12	75.1	ORF_isotig12422	gi 731346801 ref XP_010684647.1					
41.77	79	46	0	202	438	75	153	8e-12
75.1								
ORF_isotig12422	gi 983147566 gb KWU45526.1	39.39		99	57	2		
196	492	115	210	9e-12	73.2	ORF_isotig12422		
gi 731346799 ref XP_010684646.1	41.77	79	46	0	202	438	129	
207	9e-12	75.1	ORF_isotig12422	gi 490267384 ref WP_004163951.1				
29.45	146	91	2	28	438	88	230	1e-11
72.8								
ORF_isotig12422	gi 727603973 ref XP_010475030.1	37.78		90	56			
0	151	420	111	200	1e-11	74.7	ORF_isotig12422	
gi 752845383 ref WP_041474394.1	29.45	146	91	2	28	438	102	
244	1e-11	73.2	ORF_isotig12422	gi 727569811 ref XP_010457419.1				
37.78	90	56	0	151	420	111	200	1e-11
74.7								
ORF_isotig12422	gi 312171048 emb CBX79307.1	29.45		146	91	2		
28	438	88	230	1e-11	72.8	ORF_isotig12422		
gi 470239215 ref XP_004351373.1	35.14	111	66	2	190	519	475	
580	1e-11	75.1	ORF_isotig12422	gi 490258075 ref WP_004155296.1				
29.45	146	91	2	28	438	102	244	1e-11
72.8								
ORF_isotig12422	gi 310766388 gb ADP11338.1	29.45		146	91	2		
28	438	115	257	1e-11	73.2	ORF_isotig12422		
gi 219111005 ref XP_002177254.1	34.75	118	65	3	145	474	108	
221	1e-11	72.4	ORF_isotig12422	gi 478729310 emb CCP05801.1				
29.45	146	91	2	28	438	102	244	1e-11
72.8								
ORF_isotig12422	gi 635369185 emb CCI42660.1	34.38		128	80	4		
145	525	125	249	1e-11	73.2	ORF_isotig12422		
gi 752845037 ref WP_041474048.1	29.45	146	91	2	28	438	102	
244	1e-11	72.8	ORF_isotig12422	gi 817516937 ref WP_046572938.1				
37.34	158	94	4	1	474	126	278	1e-11
73.2								
ORF_isotig12422	gi 490272265 ref WP_004168384.1	29.45		146	91			
2	28	438	102	244	1e-11	72.8	ORF_isotig12422	
gi 224965368 emb CAX56900.1	29.45	146	91	2	28	438	115	257
1e-11	72.8	ORF_isotig12422	gi 297848618 ref XP_002892190.1					
37.78	90	56	0	151	420	111	200	2e-11
74.3								
ORF_isotig12422	gi 15219610 ref NP_171894.1	37.78		90	56	0		
151	420	110	199	2e-11	74.3	ORF_isotig12422		
gi 397645345 gb EJK76788.1	42.22	90	47	1	241	495	150	239
2e-11	74.3	ORF_isotig12422	gi 551588562 ref XP_005778952.1					
38.38	99	58	2	178	474	103	198	2e-11
71.6								
ORF_isotig12422	gi 301112447 ref XP_002997994.1	37.00		100	60			
2	175	474	308	404	2e-11	73.9	ORF_isotig12422	
gi 729701043 emb CEJ02028.1	42.11	76	44	0	205	432	163	238
2e-11	73.9	ORF_isotig12422	gi 754352951 ref XP_011270847.1					
41.24	97	50	2	184	456	336	431	2e-11
73.9								
ORF_isotig12422	gi 562746587 ref WP_023653709.1	30.00		140	86			
2	28	420	102	238	2e-11	72.0	ORF_isotig12422	
gi 922346013 ref XP_013449921.1	36.56	93	59	0	190	468	125	
217	2e-11	73.9	ORF_isotig12422	gi 729710313 emb CEI93301.1				
42.11	76	44	0	205	432	66	141	2e-11
73.6								
ORF_isotig12422	gi 215983522 gb ACJ71777.1	37.08		89	56	0		
202	468	128	216	2e-11	71.6	ORF_isotig12422		
gi 303281270 ref XP_003059927.1	32.22	180	93	7	10	474	57	
232	2e-11	71.6	ORF_isotig12422	gi 224001048 ref XP_002290196.1				

32.69	104	62	2	145	435	122	224	3e-11	71.6				
ORF_isotig12422			gi 552916579 gb ESA01781.1	42.67	75	43	0						
208	432	124	198	3e-11	71.2	ORF_isotig12422							
gi 727146284 emb CEG71516.1	42.11	76	44	0	205	432	139	214					
3e-11	73.6	ORF_isotig12422	gi 66803020 ref XP_635353.1	33.64									
110	68	3	190	519	252	356	3e-11	73.9	ORF_isotig12422				
gi 727146285 emb CEG71517.1	42.11	76	44	0	205	432	163	238					
3e-11	73.6	ORF_isotig12422	gi 675213985 ref XP_008915877.1										
37.00	100	60	2	175	474	303	399	3e-11	73.6				
ORF_isotig12422		gi 570996289 gb ETP52930.1	37.00	100	60	2							
175	474	302	398	3e-11	73.6	ORF_isotig12422							
gi 570335908 gb ETO83856.1	37.00	100	60	2	175	474	302	398					
3e-11	73.6	ORF_isotig12422	gi 568057987 gb ETM54635.1	37.00									
100	60	2	175	474	303	399	3e-11	73.6	ORF_isotig12422				
gi 444891677 gb AGE13358.1	51.67	60	29	0	253	432	230	289					
3e-11	72.8	ORF_isotig12422	gi 570959955 gb ETP24934.1	37.00									
100	60	2	175	474	303	399	3e-11	73.2	ORF_isotig12422				
gi 330794008 ref XP_003285073.1	40.26	77	46	0	190	420	160						
236	3e-11	73.6	ORF_isotig12422	gi 388497750 gb AFK36941.1									
41.10	73	43	0	202	420	128	200	3e-11	73.6				
ORF_isotig12422		gi 122202937 sp Q2QKL5.1 PCS3_LOTJA	41.10	73									
43	0	202	420	128	200	4e-11	73.2	ORF_isotig12422					
gi 147719992 sp Q2TE74.2 PCS2_LOTJA	41.10	73	43	0	202	420							
128	200	4e-11	73.2	ORF_isotig12422	gi 50659119 gb AAT80341.1								
41.10	73	43	0	202	420	123	195	4e-11	73.2				
ORF_isotig12422		gi 568027478 gb ETM01432.1	35.29	102	63	2							
175	480	59	157	4e-11	70.5	ORF_isotig12422							
gi 593799054 ref XP_007162565.1	42.86	77	44	0	190	420	125						
201	4e-11	73.2	ORF_isotig12422	gi 902234844 gb KNA23610.1									
38.71	93	51	2	202	474	129	217	4e-11	73.2				
ORF_isotig12422		gi 281202856 gb EFA77058.1	37.66	77	48	0							
190	420	144	220	4e-11	71.2	ORF_isotig12422							
gi 758352771 dbj GAN05107.1	33.94	109	57	2	253	534	238	346					
4e-11	73.6	ORF_isotig12422	gi 570959958 gb ETP24937.1	35.29									
102	63	2	175	480	85	183	4e-11	70.9	ORF_isotig12422				
gi 518404331 ref WP_019574538.1	45.95	74	38	1	253	474	182						
253	5e-11	71.2	ORF_isotig12422	gi 325516458 gb ADZ24787.1									
36.67	90	57	0	151	420	111	200	5e-11	7				

45.31	64	35	0	241	432	201	264	8e-11	72.4			
ORF_isotig12422				gi 661892612 emb CDP03770.1	35.51		107	62	2			
145	465	110	209	8e-11	72.4	ORF_isotig12422						
gi 823127939 ref XP_012440942.1				40.26		77	46	0	190	420	62	
138	8e-11		72.0	ORF_isotig12422		gi 566032216 gb ETI55127.1						
35.29	102	63	2	175	480	161	259	9e-11	71.2			
ORF_isotig12422				gi 4322421 gb AAD16046.1	39.73		73	44	0			
202	420	128	200	9e-11	72.0	ORF_isotig12422						
gi 571445940 ref XP_006576950.1				38.96		77	47	0	190	420	61	
137	9e-11		72.0	ORF_isotig12422		gi 662152692 ref WP_030100516.1						
36.71	158	95	4	1	474	126	278	9e-11	70.9			
ORF_isotig12422				gi 823127931 ref XP_012440911.1	40.26		77	46				
0	190	420	124	200	9e-11	72.0	ORF_isotig12422					
gi 908401434 ref XP_013070252.1				38.46		104	58	2	202	513	130	
227	9e-11		72.0	ORF_isotig12422		gi 763743549 gb KJB11048.1						
40.26	77	46	0	190	420	124	200	1e-10	72.0			
ORF_isotig12422				gi 734345461 gb KHN10759.1	38.96		77	47	0			
190	420	123	199	1e-10	72.0	ORF_isotig12422						
gi 965670943 dbj BAT85550.1				33.93		112	72	1	85	420	92	201
1e-10	72.0	ORF_isotig12422		gi 823127933 ref XP_012440917.1								
40.26	77	46	0	190	420	124	200	1e-10	72.0			
ORF_isotig12422				gi 325186679 emb CCA21228.1	37.04		108	65	3			
151	474	144	248	1e-10	70.9	ORF_isotig12422						
gi 675213991 ref XP_008915880.1				35.29		102	63	2	175	480	161	
259	1e-10		70.9	ORF_isotig12422		gi 950949688 ref XP_014495217.1						
33.93	112	72	1	85	420	92	201	1e-10	72.0			
ORF_isotig12422				gi 570335911 gb ETO83859.1	35.29		102	63	2			
175	480	161	259	1e-10	70.9	ORF_isotig12422						
gi 763743550 gb KJB11049.1				40.26		77	46	0	190	420	124	200
1e-10	72.0	ORF_isotig12422		gi 570996278 gb ETP52920.1	35.29							
102	63	2	175	480	161	259	1e-10	70.9	ORF_isotig12422			
gi 947119173 gb KRH67422.1				38.96		77	47	0	190	420	88	164
1e-10	71.6	ORF_isotig12422		gi 356505041 ref XP_003521301.1								
38.96	77	47	0	190	420	123	199	1e-10	71.6			
ORF_isotig12422				gi 167537519 ref XP_001750428.1	31.69		142	95				
2	37	456	105	246	1e-10	70.1	ORF_isotig12422					
gi 727647674 ref XP_010494405.1				37.08		89	56	0	202	468	128	
216	1e-10		71.6	ORF_isotig12422		gi 675852508 ref XP_009011814.1						
34.43	122	70	3	208	561	173	288	1e-10	71.6			
ORF_isotig12422				gi 674249977 gb KFK42742.1	35.56		90	58	0			
151	420	111	200	1e-10	71.6	ORF_isotig12422						
gi 402586327 gb EJW80265.1				39.24		79	48	0	202	438	68	146
1e-10	69.3	ORF_isotig12422		gi 567185115 ref XP_006403146.1								
38.20	89	55	0	202	468	128	216	1e-10	71.6			
ORF_isotig12422				gi 449443293 ref XP_004139414.1	35.71		98	59				
2	202	495	53	146	1e-10	68.2	ORF_isotig12422					
gi 551561391 ref XP_005767549.1				40.00		100	60	0	211	510	140	
239	1e-10		69.7	ORF_isotig12422		gi 290980155 ref XP_002672798.1						
33.90	118	67	3	145	474	153	267	1e-10	71.6			
ORF_isotig12422				gi 284466089 gb ACL00594.3	39.73		73	44	0			
202	420	128	200	2e-10	71.2	ORF_isotig12422						
gi 549049394 emb CCX11465.1				35.51		107	65	3	145	462	127	230
2e-10	71.2	ORF_isotig12422		gi 21104518 dbj BAB93120.1	34.26							
108	70	1	145	468	110	216	2e-10	71.2	ORF_isotig12422			
gi 694541843 ref XP_009497159.1				37.50		96	59	1	145	432	48	
142	2e-10		71.2	ORF_isotig12422		gi 493073505 ref WP_006122095.1						
31.52	92	63	0	145	420	148	239	2e-10	69.7			

ORF_isotig12422	gi 814540591 emb CEQ42050.1	37.80	82	51	0
256	501	156	237	2e-10	71.2 ORF_isotig12422
gi 748749862 ref WP_040007977.1	44.59	74	39	2	253 474 171
242	2e-10	69.7	ORF_isotig12422	gi 870854053 gb KMT05872.1	
41.10	73	43	0	202 420 104 176	2e-10 71.2
ORF_isotig12422	gi 870854052 gb KMT05871.1	41.10	73	43	0
202	420	129	201	2e-10	71.2 ORF_isotig12422
gi 870854054 gb KMT05873.1	41.10	73	43	0	202 420 145 217
2e-10	71.2	ORF_isotig12422	gi 970651949 gb KUF90697.1	35.29	
102	63	2	175	480	161 259 2e-10 70.5 ORF_isotig12422
gi 902234842 gb KNA23608.1	41.10	73	43	0	202 420 129 201
2e-10	71.2	ORF_isotig12422	gi 588263371 ref XP_006960829.1		
36.36	77	49	0	202 432 160 236	2e-10 70.1
ORF_isotig12422	gi 15240084 ref NP_199220.1	39.73	73	44	0
202	420	128	200	2e-10	71.2 ORF_isotig12422
gi 18254401 gb AAL66747.1 AF461180_1	39.73	73	44	0	202 420
128	200	2e-10	71.2	ORF_isotig12422	
gi 694541835 ref XP_009497157.1	37.50	96	59	1	145 432 315
409	2e-10	71.6	ORF_isotig12422		
gi 7229390 gb AAF42805.1 AF162689_1	39.73	73	44	0	202 420
128	200	2e-10	71.2	ORF_isotig12422	gi 870854055 gb KMT05874.1
41.10	73	43	0	202 420 129 201	2e-10 71.2
ORF_isotig12422	gi 4768281 gb AAD29446.1 AF085231_2	37.08	89		
56	0	202	468	145 233	2e-10 71.2 ORF_isotig12422
gi 729349445 ref XP_010542900.1	35.96	89	57	0	202 468 128
216	2e-10	71.2	ORF_isotig12422	gi 384491545 gb EIE82741.1	
29.94	157	97	2	208 639 134 290	2e-10 70.9
ORF_isotig12422	gi 671683790 emb CDS12824.1	41.56	77	45	0
202	432	206	282	2e-10	71.2 ORF_isotig12422
gi 731346805 ref XP_010684649.1	41.10	73	43	0	202 420 181
253	2e-10	71.2	ORF_isotig12422	gi 731346807 ref XP_010684650.1	
41.10	73	43	0	202 420 181 253	2e-10 71.2
ORF_isotig12422	gi 13928024 emb CAC37692.1	36.26	91	56	1
202	474	128	216	2e-10	70.9 ORF_isotig12422
gi 470508086 ref XP_004367696.1	40.00	75	45	0	196 420 108
182	2e-10	70.9	ORF_isotig12422	gi 902234843 gb KNA23609.1	
41.10	73	43	0	202 420 129 201	2e-10 70.9
ORF_isotig12422	gi 685326730 ref XP_009101739.1	36.26	91	56	
1	202	474	128	216	2e-10 70.9 ORF_isotig12422
gi 443301362 gb AGC82138.1	39.73	73	44	0	202 420 129 201
2e-10	70.9	ORF_isotig12422	gi 520899875 ref WP_020322885.1		
31.31	99	68	0	145 441 148 246	2e-10 69.3
ORF_isotig12422	gi 731346810 ref XP_010684651.1	41.10	73	43	
0	202	420	120	192	2e-10 70.9 ORF_isotig12422
gi 731346803 ref XP_010684648.1	41.10	73	43	0	202 420 181
253	2e-10	70.9	ORF_isotig12422	gi 301128435 gb ADK61091.1	
41.10	73	43	0	202 420 128 200	2e-10 70.9
ORF_isotig12422	gi 297794999 ref XP_002865384.1	39.73	73	44	
0	202	420	128	200	3e-10 70.9 ORF_isotig12422
gi 923860545 ref XP_013706438.1	36.26	91	56	1	202 474 185
273	3e-10	70.9	ORF_isotig12422	gi 565432669 ref XP_006280379.1	
34.86	109	68	2	145 468 110 216	3e-10 70.5
ORF_isotig12422	gi 170594708 ref XP_001902100.1	39.24	79	48	
0	202	438	144	222	3e-10 70.5 ORF_isotig12422
gi 702451810 ref XP_010025777.1	39.19	74	45	0	199 420 144
217	3e-10	68.9	ORF_isotig12422	gi 284468345 gb ABW98498.2	
39.73	73	44	0	202 420 128 200	3e-10 70.5

ORF_isotig12422	gi 511010013 gb EPB91244.1	45.00	60	33	0
253 432 233 292 3e-10	70.9 ORF_isotig12422				
gi 301112441 ref XP_002997991.1	37.80	82	50	1	175 420 161
241 3e-10	69.7 ORF_isotig12422	gi 928498313 gb KPD18813.1			
35.29	153 95 3 31 489 131 279 3e-10	69.3			
ORF_isotig12422	gi 672829171 gb KFH74060.1	38.53	109	56	4
190 504 488 589 3e-10	70.5 ORF_isotig12422				
gi 629091683 gb KCW57678.1	39.19	74 45 0	199	420	127 200
3e-10	68.9 ORF_isotig12422	gi 567155800 ref XP_006418181.1			
35.56	90 58 0 151 420 75 164 4e-10	70.1			
ORF_isotig12422	gi 494729398 ref WP_007465264.1	44.00	75	42	
0 208 432 171 245 4e-10	68.9 ORF_isotig12422				
gi 296417443 ref XP_002838367.1	33.03	109 71 2	145	468	110
217 4e-10	70.1 ORF_isotig12422	gi 567155803 ref XP_006418182.1			
35.56	90 58 0 151 420 111 200 4e-10	70.1			
ORF_isotig12422	gi 727621889 ref XP_010481782.1	37.08	89	56	
0 202 468 128 216 4e-10	70.1 ORF_isotig12422				
gi 21104516 dbj BAB93119.1	33.02	106 71 0	151	468	111 216
4e-10	70.1 ORF_isotig12422	gi 219129430 ref XP_002184892.1			
41.56	77 45 0 190 420 594 670 4e-10	70.5			
ORF_isotig12422	gi 727535400 ref XP_010441944.1	37.08	89	56	
0 202 468 128 216 4e-10	70.1 ORF_isotig12422				
gi 528895786 gb EPZ35810.1	35.85	106 65 2	145	462	120 222
4e-10	70.1 ORF_isotig12422	gi 502110105 ref XP_004493800.1			
42.25	71 41 0 208 420 129 199 4e-10	70.1			
ORF_isotig12422	gi 743862914 ref XP_010943667.1	39.73	73	44	
0 202 420 42 114 4e-10	70.1 ORF_isotig12422				
gi 410994522 gb AFV96147.1	34.83	89 58 0	202	468	43 131
4e-10	66.6 ORF_isotig12422	gi 747088495 ref XP_011091839.1			
38.96	77 47 0 190 420 124 200 5e-10	70.1			
ORF_isotig12422	gi 695439924 ref XP_009532833.1	36.94	111	66	
3 145 474 310 417 5e-10	70.1 ORF_isotig12422				
gi 515351240 ref WP_016863694.1	36.36	99 60 2	175	468	120
216 5e-10	69.7 ORF_isotig12422	gi 302789570 ref XP_002976553.1			
37.65	85 53 0 202 456 128 212 5e-10	70.1			
ORF_isotig12422	gi 545355774 ref XP_005643662.1	32.69	104	69	
1 145 456 83 185 5e-10	67.4 ORF_isotig12422				
gi 390364326 ref XP_003730583.1	38.06	134 74 5	46	435	76
204 5e-10	67.8 ORF_isotig12422	gi 743862910 ref XP_010943666.1			
39.73	73 44 0 202 420 66 138 5e-10	69.7			
ORF_isotig12422	gi 659070963 ref XP_008457484.1	38.36	73	45	
0 202 420 233 305 5e-10	69.3 ORF_isotig12422				
gi 19114464 ref NP_593552.1	28.70	108 76 1	145	468	145 251
6e-10	69.7 ORF_isotig12422	gi 685358931 ref XP_009114083.1			
39.73	73 44 0 202 420 156 228 6e-10	69.7			
ORF_isotig12422	gi 748629167 ref WP_039887344.1	33.10	142	91	
2 49 474 97 234 6e-10	67.8 ORF_isotig12422				
gi 674920166 emb CDY13130.1	39.73	73 44 0	202	420	110 182
6e-10	69.7 ORF_isotig12422	gi 302782732 ref XP_002973139.1			
37.65	85 53 0 202 456 128 212 6e-10	69.7			
ORF_isotig12422	gi 758368534 emb CEP09044.1	45.00	60	33	0
253 432 234 293 6e-10	69.7 ORF_isotig12422				
gi 901821666 gb KMZ73511.1	31.52	92 62 1	145	420	110 200
6e-10	69.7 ORF_isotig12422	gi 661175740 emb CDH61092.1	40.26		
77 46 0 202 432 206 282 7e-10	69.7 ORF_isotig12422				
gi 743862902 ref XP_010943664.1	39.73	73 44 0	202	420	128
200 7e-10	69.7 ORF_isotig12422	gi 18958247 dbj BAB85602.1			

35.16	91	57	1	202	474	128	216	7e-10	69.3			
ORF_isotig12422				gi 472585716 gb EMS23267.1	34.06				138	64	4	
19	432	107	217	7e-10	69.3	ORF_isotig12422						
gi 567902608 ref XP_006443792.1				39.73		73	44	0	202	420	128	
200	7e-10			69.3	ORF_isotig12422		gi 567902610 ref XP_006443793.1					
39.73		73	44	0	202	420	128	200	7e-10	69.3		
ORF_isotig12422				gi 674877470 emb CDY54482.1	35.16				91	57	1	
202	474	128	216	7e-10	69.3	ORF_isotig12422						
gi 207100019 emb CAK24968.2	35.16			91	57	1		202	474	128	216	
7e-10	69.3	ORF_isotig12422		gi 672183480 ref XP_008812020.1								
39.73		73	44	0	202	420	42	114	8e-10	69.3		
ORF_isotig12422				gi 641841508 gb KDO60420.1	39.73		73	44	0			
202	420	128	200	8e-10	69.3	ORF_isotig12422						
gi 732757650 gb KHJ79087.1	28.97			107	73	1		202	513	7	113	
8e-10	65.5	ORF_isotig12422		gi 1005176715 gb KYF56577.1	44.59							
74	37	1	253	474	150	219	8e-10	68.6	ORF_isotig12422			
gi 659071562 ref XP_008460630.1	39.02			82	44	1		202	429	235		
316	8e-10	68.6	ORF_isotig12422		gi 333458192 gb EGK86811.1							
33.10		142	91	2	49	474	127	264	8e-10	68.2		
ORF_isotig12422				gi 255577926 ref XP_002529835.1	33.33				90	60		
0	151	420	111	200	8e-10	69.3	ORF_isotig12422					
gi 694415952 ref XP_009336118.1	38.20			89	55	0		202	468	128		
216	8e-10	69.3	ORF_isotig12422		gi 720020389 ref XP_010262388.1							
36.36		77	49	0	190	420	88	164	9e-10	69.3		
ORF_isotig12422				gi 694415950 ref XP_009336117.1	38.20				89	55		
0	202	468	140	228	9e-10	69.3	ORF_isotig12422					
gi 720020382 ref XP_010262386.1	36.36			77	49	0		190	420	124		
200	9e-10	69.3	ORF_isotig12422		gi 723802071 ref NP_001289777.1							
36.36		77	49	0	190	420	124	200	9e-10	69.3		
ORF_isotig12422				gi 915281149 ref XP_013305271.1	29.52				105	73		
1	145	456	117	221	9e-10	68.2	ORF_isotig12422					
gi 255964729 gb ACU44656.1	34.44			90	59	0		151	420	111	200	
9e-10	69.3	ORF_isotig12422		gi 923847316 ref XP_013702842.1								
33.02		106	71	0	151	468	111	216	9e-10	68.6		
ORF_isotig12422				gi 672183478 ref XP_008812019.1	39.73		73	44				
0	202	420	128	200	9e-10	69.3	ORF_isotig12422					
gi 720020386 ref XP_010262387.1	36.36			77	49	0		190	420	124		
200	9e-10	69.3	ORF_isotig12422		gi 46949222 gb AAT07467.1							
38.36		73	45	0	202	420	128	200	9e-10	68.9		
ORF_isotig12422				gi 760447765 ref XP_011400951.1	34.44				90	59		
0	151	420	131	220	9e-10	68.9	ORF_isotig12422					
gi 363413051 gb AEW23125.1	38.36			73	45	0		202	420	128	200	
9e-10	68.9	ORF_isotig12422		gi 1000945507 ref XP_015581239.1								
33.33		90	60	0	151	420	111	200	9e-10	68.9		
ORF_isotig12422				gi 410994560 gb AFV96166.1	38.81				67	41	0	
202	402	103	169	1e-09	66.2	ORF_isotig12422						
gi 923762135 ref XP_013677346.1	35.16			91	57	1		202	474	184		
272	1e-09	68.9	ORF_isotig12422		gi 922533447 ref XP_013598130.1							
35.16		91	57	1	202	474	184	272	1e-09	68.9		
ORF_isotig12422				gi 220683835 gb ACL80669.1	39.73		73	44	0			
202	420	128	200	1e-09	68.9	ORF_isotig12422						
gi 676382929 ref XP_009034743.1	42.47			73	42	0		214	432	142		
214	1e-09	66.6	ORF_isotig12422		gi 923885023 ref XP_013714028.1							
39.73		73	44	0	202	420	128	200	1e-09	68.6		
ORF_isotig12422				gi 764527045 ref XP_011458022.1	35.56				90	58		
0	199	468	131	220	1e-09	68.9	ORF_isotig12422					
gi 922563714 ref XP_013610130.1	39.73			73	44	0		202	420	130		

[illegible]

gi 674936482 emb CDX96939.1	35.96	89	57	0	202	468	128	216
2e-09	67.8	ORF_isotig12422	gi 922867416 gb K0034056.1	33.63				
113	73	1	100	432	125	237	2e-09	66.6
ORF_isotig12422			gi 551535405 ref XP_005756498.1	49.18	61	31	0	250
287	2e-09	67.4	ORF_isotig12422	gi 567902606 ref XP_006443791.1				432
38.36	73	45	0	202	420	128	200	2e-09
ORF_isotig12422			gi 676493293 ref XP_009066193.1	35.06				77
0	190	420	119	195	3e-09	65.9	ORF_isotig12422	50
gi 42742267 gb AAS45236.1	34.83	89	58	0	202	468	128	216
3e-09	67.8	ORF_isotig12422	gi 312072379 ref XP_003139039.1					
36.84	76	48	0	211	438	140	215	3e-09
ORF_isotig12422			gi 407312484 gb AFU06381.1	32.26		124	75	3
145	513	110	225	3e-09	67.8	ORF_isotig12422		
gi 657961548 ref XP_008372367.1	37.66	77	48	0	202	432	128	
204	3e-09	67.8	ORF_isotig12422	gi 536747464 gb ERB64445.1				
34.62	78	51	0	187	420	175	252	3e-09
ORF_isotig12422			gi 473798304 gb EMS46513.1	32.41		108	67	1
190	513	38	139	3e-09	67.8	ORF_isotig12422		
gi 573946675 ref XP_006655670.1	31.40	121	77	1	151	513	113	
227	3e-09	67.8	ORF_isotig12422	gi 740321886 ref WP_038159030.1				
34.62	78	51	0	187	420	175	252	3e-09
ORF_isotig12422			gi 970053982 ref XP_015088621.1	33.33		90	60	
0	151	420	111	200	3e-09	67.8	ORF_isotig12422	
gi 757790478 ref WP_043008233.1	34.62	78	51	0	187	420	175	
252	3e-09	66.6	ORF_isotig12422	gi 529161423 gb AGS56990.1				
37.08	89	56	0	202	468	128	216	3e-09
ORF_isotig12422			gi 565364522 ref XP_006348973.1	36.36		77	49	
0	190	420	69	145	3e-09	67.4	ORF_isotig12422	
gi 28569702 emb CAD68108.1	33.33	90	60	0	151	420	75	164
3e-09	67.8	ORF_isotig12422	gi 972777756 ref NP_001305597.1					
33.33	90	60	0	151	420	75	164	3e-09
ORF_isotig12422			gi 444891679 gb AGE13359.1	41.56		77	45	0
202	432	205	281	3e-09	67.4	ORF_isotig12422		
gi 313483709 gb ADR51683.1	34.07	91	58	1	202	474	92	180
3e-09	67.8	ORF_isotig12422	gi 313483707 gb ADR51682.1	34.07				
91	58	1	202	474	92	180	3e-09	67.8
ORF_isotig12422			gi 567185112 ref XP_006403145.1	35.96	89	57	0	202
216	3e-09	67.8	ORF_isotig12422	gi 568215457 ref NP_001275308.1				468
33.33	90	60	0	151	420	111	200	3e-09
ORF_isotig12422			gi 373405317 gb AEY68568.1	36.26		91	56	1
202	474	128	216	3e-09	67.8	ORF_isotig12422		
gi 930111143 gb KPH26997.1	34.62	78	51	0	187	420	175	252
3e-09	66.2	ORF_isotig12422	gi 313483702 gb ADR51680.1	34.07				
91	58	1	202	474	92	180	3e-09	67.4
ORF_isotig12422			gi 740725877 ref WP_038511163.1	34.62	78	51	0	187
252	3e-09	66.2	ORF_isotig12422	gi 460404089 ref XP_004247517.1				420
33.33	90	60	0	151	420	111	200	3e-0



ORF_isotig12422	gi 659070971 ref XP_008457528.1	33.33	93	61
1	142 420 222 313 3e-09	67.4	ORF_isotig12422	
gi 743942813 ref XP_011015906.1	39.73	73	44	0 202 420 128
200 3e-09	67.4	ORF_isotig12422	gi 685369694 ref XP_009118681.1	
33.02	106 71 0 151 468 111 216 4e-09	67.4		
ORF_isotig12422	gi 674900498 emb CDY32521.1	33.02	106	71 0
151 468 111 216 4e-09	67.4	ORF_isotig12422		
gi 743869433 ref XP_010905811.1	35.37	82	51 1 175 420 35	
114 4e-09	67.0	ORF_isotig12422	gi 923703471 ref XP_013660298.1	
33.02	106 71 0 151 468 111 216 4e-09	67.4		
ORF_isotig12422	gi 166798217 gb ABY89660.1	34.44	90	59 0
151 420 111 200 4e-09	67.4	ORF_isotig12422		
gi 590715534 ref XP_007050223.1	33.33	90	60 0 151 420 113	
202 4e-09	67.4	ORF_isotig12422	gi 218197390 gb EEC79817.1	
32.41	108 67 1 190 513 38 139 4e-09	67.0		
ORF_isotig12422	gi 747072284 ref XP_011083048.1	36.71	79	50
0 202 438 128 206 4e-09	67.4	ORF_isotig12422		
gi 939611289 gb KPV73315.1	47.46	59	31 0 256 432 161 219	
4e-09	67.0	ORF_isotig12422	gi 695064415 ref XP_009420754.1	
36.99	73 46 0 202 420 128 200 4e-09	67.0		
ORF_isotig12422	gi 685369692 ref XP_009118680.1	33.02	106	71
0 151 468 111 216 4e-09	67.0			

ORF_comp2951_c3_seq1	gi 501733845 ref WP_012629148.1	39.31	145
81 2 115 549 44 181 5e-23	103	ORF_comp2951_c3_seq1	
gi 493029431 ref WP_006099075.1	38.73	142	80 2 121 546 47
181 1e-22	102	ORF_comp2951_c3_seq1	gi 740234212 ref WP_038075400.1
41.10	146 77 4 115 546 36 174 2e-22	101	
ORF_comp2951_c3_seq1	gi 648291774 ref WP_026072923.1	39.44	142
78 2 121 546 46 179 4e-22	100	ORF_comp2951_c3_seq1	
gi 501377640 ref WP_012409206.1	37.93	145	82 4 115 546 41
178 9e-22	100	ORF_comp2951_c3_seq1	gi 505020123 ref WP_015207225.1
37.93	145 82 3 115 546 44 181 9e-21	97.4	
ORF_comp2951_c3_seq1	gi 738534441 ref WP_036482850.1	35.42	144
90 2 118 546 46 187 1e-20	97.1	ORF_comp2951_c3_seq1	
gi 658419581 ref WP_029646255.1	33.80	142	87 2 121 546 47
181 1e-20	97.1	ORF_comp2951_c3_seq1	
gi 522054471 ref WP_020565680.1	35.81	148	76 3 121 546 54
188 2e-20	96.7	ORF_comp2951_c3_seq1	
gi 499931312 ref WP_011612046.1	37.93	145	80 3 115 546 45
180 2e-20	96.3	ORF_comp2951_c3_seq1	
gi 515385832 ref WP_016877723.1	38.19	144	79 4 121 546 46
181 3e-20	95.5	ORF_comp2951_c3_seq1	gi 689678301 emb CEF48871.1
37.93	145 77 3 121 546 40 174 4e-20	95.5	
ORF_comp2951_c3_seq1	gi 656719681 gb KEF42411.1	38.78	147 80
4 112 546 22 160 4e-20	95.1	ORF_comp2951_c3_seq1	
gi 981461217 ref WP_059669889.1	37.93	145	77 3 121 546 40
174 6e-20	94.7	ORF_comp2951_c3_seq1	
gi 763312982 ref WP_044171432.1	37.32	142	82 2 121 546 38
172 7e-20	94.7	ORF_comp2951_c3_seq1	
gi 826018244 ref WP_047165229.1	37.93	145	80 3 115 546 33
168 1e-19	94.4	ORF_comp2951_c3_seq1	
gi 504925739 ref WP_015112841.1	38.46	143	80 3 121 546 46
181 2e-19	94.0	ORF_comp2951_c3_seq1	
gi 516956783 ref WP_018183684.1	35.95	153	88 4 121 579 39
181 2e-19	93.6	ORF_comp2951_c3_seq1	

gi 503088795 ref WP_013323642.1	37.50	144	82	3	121	549	47
183 3e-19	93.2 ORF_comp2951_c3_seq1						
gi 748165414 ref WP_039738990.1	36.36	143	83	4	121	546	46
181 4e-19	92.4 ORF_comp2951_c3_seq1						
gi 515356232 ref WP_016865554.1	37.24	145	83	4	115	546	43
180 6e-19	92.0 ORF_comp2951_c3_seq1						
gi 941312503 ref WP_055116231.1	36.05	147	85	3	112	546	35
174 6e-19	92.0 ORF_comp2951_c3_seq1						
gi 751277775 ref WP_040986624.1	36.11	144	85	3	115	546	36
172 7e-19	92.0 ORF_comp2951_c3_seq1						
gi 751301901 ref WP_041010414.1	36.11	144	85	3	115	546	36
172 7e-19	92.0 ORF_comp2951_c3_seq1						
gi 751301199 ref WP_041009719.1	36.11	144	85	3	115	546	36
172 8e-19	91.7 ORF_comp2951_c3_seq1	gi 685089209 emb CDX24205.1					
36.11	144 85 3 115 546 36	172 8e-19	91.7				
ORF_comp2951_c3_seq1	gi 751264288 ref WP_040974001.1	36.11	144				
85 3 115 546 38	174 8e-19	91.7 ORF_comp2951_c3_seq1					
gi 737933697 ref WP_035898371.1	36.05	147	85	3	112	546	35
174 9e-19	91.7 ORF_comp2951_c3_seq1						
gi 648412694 ref WP_026104445.1	35.66	143	85	3	121	546	46
182 1e-18	91.7 ORF_comp2951_c3_seq1						
gi 917776332 ref WP_052290273.1	36.11	144	84	3	118	546	33
169 2e-18	90.9 ORF_comp2951_c3_seq1	gi 746961585 gb KIF19114.1					
36.11	144 84 3 118 546 42	178 2e-18	90.9				
ORF_comp2951_c3_seq1	gi 927917924 emb CDX53272.1	36.11	144	85			
3 115 546 36	172 2e-18	90.9 ORF_comp2951_c3_seq1					
gi 652338179 ref WP_026735392.1	36.55	145	84	4	115	546	44
181 2e-18	90.9 ORF_comp2951_c3_seq1						
gi 493165631 ref WP_006170876.1	39.04	146	81	3	112	546	69
207 3e-18	90.5 ORF_comp2951_c3_seq1						
gi 504945810 ref WP_015132912.1	35.92	142	83	3	121	546	47
180 4e-18	89.7 ORF_comp2951_c3_seq1						
gi 505018197 ref WP_015205299.1	37.76	143	81	3	121	546	46
181 4e-18	89.7 ORF_comp2951_c3_seq1	gi 814343774 emb CPH68122.1					
36.62	142 83 3 121 546 38	172 5e-18	89.7				
ORF_comp2951_c3_seq1	gi 953510847 dbj BAT53535.1	35.62	146	84			
5 115 546 36	173 5e-18	89.4 ORF_comp2951_c3_seq1					
gi 916294818 ref WP_051029864.1	36.99	146	82	4	115	546	42
179 5e-18	89.4 ORF_comp2951_c3_seq1						
gi 737172391 ref WP_035158595.1	36.36	143	83	4	121	546	43
178 5e-18	89.4 ORF_comp2951_c3_seq1						
gi 821561551 ref WP_046868556.1	37.32	142	82	3	121	546	41
175 6e-18	89.4 ORF_comp2951_c3_seq1						
gi 497576287 ref WP_009890471.1	37.06	143	81	4	121	546	40
174 6e-18	89.4 ORF_comp2951_c3_seq1	gi 584107904 gb AHI79522.1					
37.06	143 81 4 121 546 38	172 6e-18	89.4				
ORF_comp2951_c3_seq1	gi 773064690 gb AJY42591.1	35.81	148	82			
3 121 555 30	167 8e-18	89.0 ORF_comp2951_c3_seq1					
gi 495550750 ref WP_008275329.1	37.06	143	82	3	121	546	42
177 8e-18	89.0 ORF_comp2951_c3_seq1						
gi 499442921 ref WP_011130385.1	38.89	144	78	5	121	546	52
187 8e-18	89.0 ORF_comp2951_c3_seq1						
gi 492895672 ref WP_006026078.1	35.81	148	82	3	121	555	40
177 9e-18	89.0 ORF_comp2951_c3_seq1						
gi 756829253 ref WP_042646352.1	35.42	144	86	3	115	546	36
172 9e-18	89.0 ORF_comp2951_c3_seq1						
gi 981645748 ref WP_059844321.1	37.06	143	81	4	121	546	40

[illegible]

[illegible]

ORF_comp2951_c3_seq1	gi 157806771 gb EDO83941.1	35.86	145	80
3	121 546 158 292 4e-16	85.1	ORF_comp2951_c3_seq1	
gi 225927831 gb EEH23872.1	35.86	145	80	3
4e-16	85.5	ORF_comp2951_c3_seq1	gi 134251017 gb EBA51096.1	
35.86	145 80 3	121 546 166	300 4e-16	85.5
ORF_comp2951_c3_seq1	gi 237502694 gb ACQ95012.1	35.86	145	80
3	121 546 144 278 4e-16	85.1	ORF_comp2951_c3_seq1	
gi 714902087 gb KGX30928.1	35.86	145	80	3
4e-16	85.1	ORF_comp2951_c3_seq1	gi 640368848 ref WP_024880427.1	
35.86	145 78 3	121 546 43	175 5e-16	84.0
ORF_comp2951_c3_seq1	gi 85544315 pdb 2BTW B	33.33	144	86
121 546 52	187 6e-16	84.0	ORF_comp2951_c3_seq1	
gi 85544314 pdb 2BTW A	33.33	144	86	5
16	83.6	ORF_comp2951_c3_seq1	gi 817516937 ref WP_046572938.1	
35.92	142 84 2	112 537 80	214 6e-16	84.0
ORF_comp2951_c3_seq1	gi 751564302 ref WP_041033204.1	33.57	143	
87 3	121 546 46	181 7e-16	83.2	ORF_comp2951_c3_seq1
gi 741007236 ref WP_038789710.1	35.17	145	81	3
174 8e-16	83.2	ORF_comp2951_c3_seq1	gi 741019975 ref WP_038801977.1	
gi 741019975 ref WP_038801977.1	35.86	145	80	3
173 8e-16	83.2	ORF_comp2951_c3_seq1	gi 772972989 gb AJX33047.1	
35.86	145 80 3	121 546 38	172 9e-16	83.2
ORF_comp2951_c3_seq1	gi 740943421 ref WP_038728427.1	35.17	145	
81 3	121 546 40	174 9e-16	83.2	ORF_comp2951_c3_seq1
gi 785594494 ref WP_045719475.1	35.17	145	81	3
174 9e-16	83.2	ORF_comp2951_c3_seq1	gi 806494169 gb KKC16091.1	
35.17	145 81 3	121 546 30	164 9e-16	82.8
ORF_comp2951_c3_seq1	gi 938273386 gb KPQ08481.1	35.21	142	85
3	121 546 32 166 1e-15	82.8	ORF_comp2951_c3_seq1	
gi 123963072 gb ABM77828.1	36.11	144	82	5
1e-15	83.2	ORF_comp2951_c3_seq1	gi 496440563 ref WP_009149408.1	
34.51	142 85 2	121 546 60	193 1e-15	82.8
ORF_comp2951_c3_seq1	gi 740986897 ref WP_038770440.1	35.17	145	
81 3	121 546 40	174 1e-15	82.4	ORF_comp2951_c3_seq1
gi 568795567 gb AHE26011.1	35.17	145	81	3
1e-15	82.4	ORF_comp2951_c3_seq1	gi 427346779 gb AFY29492.1	
35.62	146 86 3	112 546 25	163 2e-15	82.4
ORF_comp2951_c3_seq1	gi 652914138 ref WP_027168079.1	33.33	144	
89 3	115 546 36	172 2e-15	82.4	ORF_comp2951_c3_seq1
gi 686936961 gb KGD15055.1	37.04	135	72	3
2e-15	81.6	ORF_comp2951_c3_seq1	gi 740967447 ref WP_038751917.1	
35.17	145 81 3	121 546 40	174 2e-15	82.0
ORF_comp2951_c3_seq1	gi 759607566 ref WP_043325766.1	35.62	146	
86 3	112 546 36	174 2e-15	82.0	ORF_comp2951_c3_seq1
gi 738342341 ref WP_036294947.1	35.17	145	79	3
180 3e-15	81.6	ORF_comp2951_c3_seq1	gi 737586499 ref WP_035557439.1	
gi 737586499 ref WP_035557439.1	35.92	142	84	2
227 3e-15	82.0	ORF_comp2951_c3_seq1	gi 714591842 gb KGW22095.1	
35.17	145 81 3	121 546 30	164 3e-15	81.6
ORF_comp2951_c3_seq1	gi 501395618 ref WP_012427184.1	37.32	142	
82 3	112 537 81	215 3e-15	82.0	ORF_comp2951_c3_seq1
gi 651604058 ref WP_026598946.1	35.17	145	79	3
170 4e-15	81.3	ORF_comp2951_c3_seq1	gi 814476642 emb CPI32267.1	
35.17	145 81 3	121 546 38	172 5e-15	80.9
ORF_comp2951_c3_seq1	gi 648499655 ref WP_026191406.1	35.17	145	
79 3	121 546 43	175 5e-15	80.9	ORF_comp2951_c3_seq1
gi 917319449 ref WP_051926161.1	35.71	140	82	3
		121	537	46

[illegible]

gi 213409680 ref XP_002175610.1	30.34	145	85	5	121	546	59
190 9e-08	61.2 ORF_comp2951_c3_seq1						
gi 966405784 ref WP_058449190.1	32.94	85	49	2	157	408	9
86 9e-08	59.3 ORF_comp2951_c3_seq1						
gi 503689125 ref WP_013923201.1	36.11	108	58	5	121	429	37
138 1e-07	58.5 ORF_comp2951_c3_seq1						
gi 490267384 ref WP_004163951.1	29.27	164	93	5	121	546	17
179 2e-07	59.3 ORF_comp2951_c3_seq1						
gi 918692141 ref WP_052561880.1	30.00	140	83	3	130	549	40
164 2e-07	58.9 ORF_comp2951_c3_seq1						
gi 490258075 ref WP_004155296.1	29.27	164	93	5	121	546	31
193 2e-07	59.3 ORF_comp2951_c3_seq1						
gi 831778908 ref XP_012755554.1	25.38	130	82	4	160	546	85
200 4e-07	59.3 ORF_comp2951_c3_seq1	gi 925230922 gb KOS75550.1					
33.94	109 59 3 229 546 1 99	4e-07	57.0				
ORF_comp2951_c3_seq1	gi 813208513 dbj GAO50417.1	29.32	133	77			
4 154 546 146 263 5e-07	59.3 ORF_comp2951_c3_seq1						
gi 310766388 gb ADP11338.1	29.27	164	93	5	121	546	44 206
5e-07	58.5 ORF_comp2951_c3_seq1	gi 752845383 ref WP_041474394.1					
29.27	164 93 5 121 546 31	193 5e-07	58.2				
ORF_comp2951_c3_seq1	gi 752845037 ref WP_041474048.1	29.27	164				
93 5 121 546 31 193 8e-07	57.8 ORF_comp2951_c3_seq1						
gi 224965368 emb CAX56900.1	29.27	164	93	5	121	546	44 206
9e-07	57.8 ORF_comp2951_c3_seq1	gi 915474789 ref WP_050814322.1					
30.84	107 65 3 229 546 1 99	1e-06	55.8				
ORF_comp2951_c3_seq1	gi 872579356 ref WP_048533013.1	27.91	129				
84 4 157 537 144 265 1e-06	57.8 ORF_comp2951_c3_seq1						
gi 497652379 ref WP_009966563.1	30.28	142	69	3	121	546	40
151 2e-06	56.2 ORF_comp2951_c3_seq1	gi 772943476 gb AJX80078.1					
30.28	142 69 3 121 546 30 141	2e-06	55.8				
ORF_comp2951_c3_seq1	gi 736937357 ref WP_034934372.1	28.05	164				
95 5 121 546 32 194 3e-06	56.2 ORF_comp2951_c3_seq1						
gi 517208875 ref WP_018397693.1	38.38	99	47	2	121	417	50
134 3e-06	54.7 ORF_comp2951_c3_seq1						
gi 923123265 ref XP_013755101.1	28.79	132	79	5	160	546	39
158 3e-06	56.6 ORF_comp2951_c3_seq1						
gi 308510244 ref XP_003117305.1	29.01	131	78	5	160	546	49
166 3e-06	56.6 ORF_comp2951_c3_seq1						
gi 737401579 ref WP_035382755.1	32.18	87	54	1	157	417	44
125 4e-06	55.5 ORF_comp2951_c3_seq1						
gi 493073505 ref WP_006122095.1	27.11	166	98	5	115	546	30
194 4e-06	55.8 ORF_comp2951_c3_seq1	gi 672825823 gb KFH70714.1					
30.30	132 76 4 154 546 146 262	5e-06	56.2				
ORF_comp2951_c3_seq1	gi 857976683 emb CEO97478.1	26.72	131	81			
3 160 546 39 156 8e-06	55.5 ORF_comp2951_c3_seq1						
gi 953489559 emb CEG46925.1	30.23	129	78	4	160	546	188 304
9e-06	55.5 ORF_comp2951_c3_seq1	gi 443690998 gb ELT92982.1					
25.00	132 87 3 154 549 40 159	1e-05	54.3				
ORF_comp2951_c3_seq1	gi 800137352 dbj GAO33130.1	29.13	103	63			
2 157 450 46 143 1e-05	54.3 ORF_comp2951_c3_seq1						
gi 159485978 ref XP_001701021.1	27.69	130	81	3	160	546	17
134 1e-05	53.1 ORF_comp2951_c3_seq1	gi 341888907 gb EGT44842.1					
28.46	130 80 3 160 546 41 158	2e-05	54.7				
ORF_comp2951_c3_seq1	gi 502816264 ref WP_013051240.1	28.14	167				
93 7 121 549 38 201 2e-05	54.3 ORF_comp2951_c3_seq1						
gi 290993528 ref XP_002679385.1	32.06	131	75	5	154	546	56
172 2e-05	53.5 ORF_comp2951_c3_seq1	gi 309365040 emb CAP23744.2					





ORF_comp2951_c3_seq2	gi 493029431 ref WP_006099075.1	40.29	206
115 3 121 735 47 245 4e-39	150 ORF_comp2951_c3_seq2		
gi 751277775 ref WP_040986624.1	40.98 205 112 4 115 723 36		
233 5e-39 150 ORF_comp2951_c3_seq2	gi 751301901 ref WP_041010414.1		
40.98 205 112 4 115 723 36 233 5e-39 150			
ORF_comp2951_c3_seq2	gi 751301199 ref WP_041009719.1	40.98 205	
112 4 115 723 36 233 7e-39	149 ORF_comp2951_c3_seq2		
gi 501377640 ref WP_012409206.1	39.81 206 115 5 115 726 41		
239 8e-39 149 ORF_comp2951_c3_seq2	gi 522054471 ref WP_020565680.1		
38.91 221 111 5 121 762 54 257 2e-38 149			
ORF_comp2951_c3_seq2	gi 515385832 ref WP_016877723.1	41.55 207	
106 6 121 726 46 242 3e-38	148 ORF_comp2951_c3_seq2		
gi 505020123 ref WP_015207225.1	41.06 207 111 5 115 726 44		
242 3e-38 148 ORF_comp2951_c3_seq2	gi 763312982 ref WP_044171432.1		
40.29 206 115 3 121 735 38 236 7e-38 147			
ORF_comp2951_c3_seq2	gi 499931312 ref WP_011612046.1	41.46 205	
109 4 115 723 45 240 7e-38	147 ORF_comp2951_c3_seq2		
gi 756829253 ref WP_042646352.1	40.49 205 113 4 115 723 36		
233 1e-37 146 ORF_comp2951_c3_seq2	gi 751264288 ref WP_040974001.1		
40.00 205 114 4 115 723 38 235 1e-37 146			
ORF_comp2951_c3_seq2	gi 685089209 emb CDX24205.1	40.00 205 114	
4 115 723 36 233 1e-37 146 ORF_comp2951_c3_seq2			
gi 689678301 emb CEF48871.1	40.78 206 107 4 121 723 40 235		
1e-37 146 ORF_comp2951_c3_seq2	gi 503088795 ref WP_013323642.1		
40.69 204 112 4 121 726 47 243 2e-37 146			
ORF_comp2951_c3_seq2	gi 981461217 ref WP_059669889.1	41.75 206	
105 4 121 723 40 235 2e-37 145 ORF_comp2951_c3_seq2			
gi 504925739 ref WP_015112841.1	40.20 204 113 4 121 726 46		
242 3e-37 145 ORF_comp2951_c3_seq2	gi 927917924 emb CDX53272.1		
40.00 205 114 4 115 723 36 233 3e-37 145			
ORF_comp2951_c3_seq2	gi 826018244 ref WP_047165229.1	41.46 205	
109 4 115 723 33 228 4e-37 145 ORF_comp2951_c3_seq2			
gi 656719681 gb KEF42411.1	43.00 207 105 6 112 720 22 219		
6e-37 144 ORF_comp2951_c3_seq2	gi 748165414 ref WP_039738990.1		
39.90 203 113 5 121 723 46 241 6e-37 144			
ORF_comp2951_c3_seq2	gi 814343774 emb CPH68122.1	42.23 206 104	
4 121 723 38 233 7e-37 144 ORF_comp2951_c3_seq2			
gi 648291774 ref WP_026072923.1	40.29 206 114 3 121 735 46		
243 7e-37 144 ORF_comp2951_c3_seq2	gi 648412694 ref WP_026104445.1		
38.65 207 119 4 121 735 46 246 8e-37 144			
ORF_comp2951_c3_seq2	gi 505018197 ref WP_015205299.1	41.18 204	
109 5 121 723 46 241 1e-36 144 ORF_comp2951_c3_seq2			
gi 484585464 gb AGK47089.1	41.26 206 106 4 121 723 30 225		
2e-36 143 ORF_comp2951_c3_seq2	gi 499442921 ref WP_011130385.1		
42.36 203 106 6 121 720 52 246 2e-36 144			
ORF_comp2951_c3_seq2	gi 754360430 ref WP_041861630.1	41.26 206	
106 4 121 723 40 235 2e-36 143 ORF_comp2951_c3_seq2			
gi 773064690 gb AJY42591.1	41.26 206 106 4 121 723 30 225		
2e-36 142 ORF_comp2951_c3_seq2	gi 981741072 ref WP_059934588.1		
41.26 206 106 4 121 723 40 235 2e-36 143			
ORF_comp2951_c3_seq2	gi 737172391 ref WP_035158595.1	39.32 206	
112 5 121 726 43 239 3e-36 143 ORF_comp2951_c3_seq2			
gi 738534441 ref WP_036482850.1	37.44 203 123 3 118 720 46		
246 3e-36 143 ORF_comp2951_c3_seq2	gi 492895672 ref WP_006026078.1		
41.26 206 106 4 121 723 40 235 3e-36 142			
ORF_comp2951_c3_seq2	gi 913438795 ref WP_050418395.1	40.39 203	
112 4 121 723 40 235 3e-36 142 ORF_comp2951_c3_seq2			

gi 501395618 ref WP_012427184.1	42.11	209	113	4	112	735	81
282 3e-36	144 ORF_comp2951_c3_seq2	gi 909830227 gb AKU13025.1					
40.39	203 112 4	121 723 45	240 3e-36	142			
ORF_comp2951_c3_seq2	gi 584107904 gb AHI79522.1	41.67	204	108			
5 121 723 38	233 4e-36	142 ORF_comp2951_c3_seq2					
gi 505013627 ref WP_015200729.1	37.98	208	120	4	118	735	45
245 4e-36	142 ORF_comp2951_c3_seq2	gi 974600092 ref WP_059190108.1					
38.86	211 120 4	115 741 36	239 4e-36	142			
ORF_comp2951_c3_seq2	gi 740984489 ref WP_038768164.1	41.75	206				
105 4 121 723 40	235 4e-36	142 ORF_comp2951_c3_seq2					
gi 497576287 ref WP_009890471.1	41.67	204	108	5	121	723	40
235 4e-36	142 ORF_comp2951_c3_seq2	gi 657935844 ref WP_029636980.1					
39.41	203 114 5	121 723 46	241 4e-36	142			
ORF_comp2951_c3_seq2	gi 497590677 ref WP_009904861.1	41.38	203				
110 4 121 723 40	235 5e-36	142 ORF_comp2951_c3_seq2					
gi 782673738 ref WP_045600604.1	41.38	203	110	4	121	723	40
235 5e-36	142 ORF_comp2951_c3_seq2	gi 493165631 ref WP_006170876.1					
43.08	195 103 3	112 693 69	256 5e-36	143			
ORF_comp2951_c3_seq2	gi 91691483 gb ABE34681.1	40.67	209	116			
3 112 735 44	245 6e-36	142 ORF_comp2951_c3_seq2					
gi 981383968 ref WP_059597148.1	41.75	206	105	4	121	723	30
225 6e-36	142 ORF_comp2951_c3_seq2	gi 953510847 dbj BAT53535.1					
38.83	206 115 6	115 723 36	233 6e-36	142			
ORF_comp2951_c3_seq2	gi 515356232 ref WP_016865554.1	39.32	206				
116 5 115 726 43	241 7e-36	142 ORF_comp2951_c3_seq2					
gi 981645748 ref WP_059844321.1	41.67	204	108	5	121	723	40
235 7e-36	142 ORF_comp2951_c3_seq2	gi 821561551 ref WP_046868556.1					
41.38	203 110 4	121 723 41	236 8e-36	141			
ORF_comp2951_c3_seq2	gi 652338179 ref WP_026735392.1	40.38	208				
111 6 115 726 44	242 1e-35	141 ORF_comp2951_c3_seq2					
gi 749301086 ref WP_040125311.1	40.67	209	116	3	112	735	94
295 1e-35	142 ORF_comp2951_c3_seq2	gi 976471317 gb KVD74754.1					
40.78	206 107 4	121 723 30	225 2e-35	140			
ORF_comp2951_c3_seq2	gi 917776332 ref WP_052290273.1	39.81	206				
111 5 118 723 33	229 2e-35	140 ORF_comp2951_c3_seq2					
gi 746961585 gb KIF19114.1	39.81	206	111	5	118	723	42
2e-35	140 ORF_comp2951_c3_seq2	gi 941312503 ref WP_055116231.1					
38.65	207 116 4	112 720 35	234 2e-35	140			
ORF_comp2951_c3_seq2	gi 981296166 ref WP_059513522.1	40.89	203				
111 4 121 723 39	234 2e-35	140 ORF_comp2951_c3_seq2					
gi 740959588 ref WP_038744210.1	41.26	206	106	4	121	723	40
235 2e-35	140 ORF_comp2951_c3_seq2	gi 714646342 gb KGW52040.1					
41.26	206 106 4	121 723 30	225 3e-35	140			
ORF_comp2951_c3_seq2	gi 656037395 ref WP_029076335.1	38.46	208				
119 4 121 738 39	239 3e-35	140 ORF_comp2951_c3_seq2					
gi 976553572 gb KVE52227.1	40.78	206	107	4	121	723	30
3e-35	140 ORF_comp2951_c3_seq2	gi 504949901 ref WP_015137003.1					
39.41	203 114 4	121 723 46	241 3e-35	140			
ORF_comp2951_c3_seq2	gi 494317175 ref WP_007177410.1	41.95	205				
111 4 112 723 90	287 3e-35	141 ORF_comp2951_c3_seq2					
gi 752524486 ref WP_041196068.1	41.26	206	106	4	121	723	40
235 3e-35	140 ORF_comp2951_c3_seq2	gi 913356948 ref WP_050376444.1					
41.26	206 106 4	121 723 40	235 3e-35	140			
ORF_comp2951_c3_seq2	gi 800785064 emb CFL25926.1	41.26	206	106			
4 121 723 38	233 4e-35	140 ORF_comp2951_c3_seq2					
gi 953804284 ref WP_058039816.1	41.26	206	106	4	121	723	40
235 4e-35	140 ORF_comp2951_c3_seq2	gi 490685647 ref WP_004550267.1					

41.26	206	106	4	121	723	38	233	4e-35	140
ORF_comp2951_c3_seq2	gi 714507186 gb KGV37859.1	41.26	206	106	4	121	723	40	235
4	121	723	30	225	4e-35	139	ORF_comp2951_c3_seq2	gi 490301683 ref WP_004197081.1	41.26
235	4e-35	140	ORF_comp2951_c3_seq2	gi 953803876 ref WP_058039691.1	41.26	206	106	4	121
41.26	206	106	4	121	723	40	235	4e-35	140
ORF_comp2951_c3_seq2	gi 490304232 ref WP_004199532.1	41.26	206	106	4	121	723	39	234
106	4	121	723	30	225	4e-35	139	ORF_comp2951_c3_seq2	gi 981251097 ref WP_059470554.1
gi 981251097 ref WP_059470554.1	41.26	206	106	4	121	723	39	234	4e-35
38.65	207	116	4	112	720	35	234	4e-35	139
ORF_comp2951_c3_seq2	gi 504970487 ref WP_015157589.1	39.71	204	112	5	121	723	40	235
112	5	121	723	40	235	6e-35	139	ORF_comp2951_c3_seq2	gi 705475049 gb KGS04910.1
gi 705475049 gb KGS04910.1	40.78	206	107	4	121	723	29	224	9e-35
9e-35	138	ORF_comp2951_c3_seq2	gi 75702893 gb ABA22569.1	38.05	205	116	6	121	726
205	116	6	121	726	38	234	1e-34	138	ORF_comp2951_c3_seq2
gi 981433289 ref WP_059643234.1	40.78	206	107	4	121	723	40	235	1e-34
235	1e-34	139	ORF_comp2951_c3_seq2	gi 740960548 ref WP_038745121.1	40.78	206	107	4	121
40.78	206	107	4	121	723	40	235	1e-34	138
ORF_comp2951_c3_seq2	gi 504945810 ref WP_015132912.1	39.13	207	117	4	121	738	47	245
117	4	121	738	47	245	1e-34	139	ORF_comp2951_c3_seq2	gi 160698859 gb EDP88829.1
gi 160698859 gb EDP88829.1	40.89	203	111	4	121	723	136	331	1e-34
1e-34	140	ORF_comp2951_c3_seq2	gi 148027253 gb EDK85274.1	40.89	203	111	4	121	723
203	111	4	121	723	140	335	1e-34	140	ORF_comp2951_c3_seq2
gi 243063500 gb EES45686.1	40.89	203	111	4	121	723	140	335	1e-34
1e-34	140	ORF_comp2951_c3_seq2	gi 785594494 ref WP_045719475.1	40.78	206	107	4	121	723
40.78	206	107	4	121	723	40	235	1e-34	138
ORF_comp2951_c3_seq2	gi 126229014 gb ABN92554.1	40.89	203	111	4	121	723	142	337
4	121	723	145	340	1e-34	141	ORF_comp2951_c3_seq2	gi 121227363 gb ABM49881.1	40.89
gi 121227363 gb ABM49881.1	40.89	203	111	4	121	723	142	337	1e-34
1e-34	140	ORF_comp2951_c3_seq2	gi 741007236 ref WP_038789710.1	40.78	206	107	4	121	723
40.78	206	107	4	121	723	40	235	1e-34	138
ORF_comp2951_c3_seq2	gi 157806771 gb EDO83941.1	40.89	203	111	4	121	723	155	350
4	121	723	158	353	1e-34	141	ORF_comp2951_c3_seq2	gi 169654349 gb EDS87042.1	40.89
gi 169654349 gb EDS87042.1	40.89	203	111	4	121	723	155	350	1e-34
1e-34	141	ORF_comp2951_c3_seq2	gi 806494169 gb KKC16091.1	40.78	206	107	4	121	723
206	107	4	121	723	30	225	1e-34	138	ORF_comp2951_c3_seq2
gi 740943421 ref WP_038728427.1	40.78	206	107	4	121	723	40	235	1e-34
235	1e-34	138	ORF_comp2951_c3_seq2	gi 237502694 gb ACQ95012.1	40.89	203	111	4	121
40.89	203	111	4	121	723	144	339	2e-34	140
ORF_comp2951_c3_seq2	gi 126219717 gb ABN83223.1	40.8							

[illegible]

ORF_comp2951_c3_seq2	gi 739030777 ref WP_036911908.1	40.39	203
110 6 121 720 52 246 1e-32	133 ORF_comp2951_c3_seq2		
gi 648499655 ref WP_026191406.1	39.81 206 107 4 121 723 43		
236 1e-32 133 ORF_comp2951_c3_seq2	gi 966518160 ref WP_058534654.1		
36.36 198 117 5 130 723 38 226 1e-32	132		
ORF_comp2951_c3_seq2	gi 938273386 gb KPQ08481.1	39.11	202 115
4 121 723 32 226 1e-32	132 ORF_comp2951_c3_seq2		
gi 85544315 pdb 2BTW B	36.59 205 119 6 121 726 52 248 2e-		
32 133 ORF_comp2951_c3_seq2	gi 85544314 pdb 2BTW A	36.59 205 119	
6 121 726 52 248 2e-32	132 ORF_comp2951_c3_seq2		
gi 697780908 gb AIT63532.1	35.41 209 128 3 115 735 22 225		
2e-32 132 ORF_comp2951_c3_seq2	gi 817516937 ref WP_046572938.1		
41.43 210 115 3 112 738 80 282 3e-32	133		
ORF_comp2951_c3_seq2	gi 738342341 ref WP_036294947.1	39.32	206
108 4 121 723 48 241 5e-32	131 ORF_comp2951_c3_seq2		
gi 207082025 gb ABS78257.2	35.41 209 128 3 115 735 43 246		
5e-32 131 ORF_comp2951_c3_seq2	gi 651604058 ref WP_026598946.1		
38.83 206 109 4 121 723 38 231 5e-32	131		
ORF_comp2951_c3_seq2	gi 737586499 ref WP_035557439.1	42.11	209
113 3 112 735 93 294 5e-32	132 ORF_comp2951_c3_seq2		
gi 165917203 gb EDR35807.1	37.23 188 113 1 115 678 81 263		
8e-32 132 ORF_comp2951_c3_seq2	gi 489705594 ref WP_003609731.1		
39.90 203 109 5 121 723 45 236 1e-31	130		
ORF_comp2951_c3_seq2	gi 917319449 ref WP_051926161.1	39.41	203
112 5 121 720 46 240 2e-31	130 ORF_comp2951_c3_seq2		
gi 427346779 gb AFY29492.1	39.02 205 116 4 112 720 25 222		
3e-31 129 ORF_comp2951_c3_seq2	gi 759607566 ref WP_043325766.1		
39.02 205 116 4 112 720 36 233 3e-31	129		
ORF_comp2951_c3_seq2	gi 662152692 ref WP_030100516.1	40.48	210
117 3 112 738 80 282 3e-31	130 ORF_comp2951_c3_seq2		
gi 654738300 ref WP_028194554.1	40.48 210 117 3 112 738 80		
282 4e-31 130 ORF_comp2951_c3_seq2	gi 797211671 ref WP_045872614.1		
38.24 204 115 5 121 723 46 241 6e-31	128		
ORF_comp2951_c3_seq2	gi 981354997 ref WP_059570263.1	40.39	203
112 4 121 723 40 235 1e-30	127 ORF_comp2951_c3_seq2		
gi 916990632 ref WP_051597344.1	35.23 193 118 3 157 735 58		
243 8e-30 125 ORF_comp2951_c3_seq2	gi 928498313 gb KPD18813.1		
39.45 218 108 4 112 738 76 278 1e-29	125		
ORF_comp2951_c3_seq2	gi 504939174 ref WP_015126276.1	36.84	209
121 4 118 738 53 252 1e-28	122 ORF_comp2951_c3_seq2		
gi 827519218 gb KLN61905.1	33.33 204 128 4 115 723 24 220		
8e-28 119 ORF_comp2951_c3_seq2	gi 500232799 ref WP_011902392.1		
34.98 203 120 6 121 720 38 231 2e-27	119		
ORF_comp2951_c3_seq2	gi 920578875 ref WP_053005650.1	33.33	204
128 4 115 723 40 236 2e-27	119 ORF_comp2951_c3_seq2		
gi 167537519 ref XP_001750428.1	38.24 204 106 5 121 702 48		
241 4e-27 118 ORF_comp2951_c3_seq2	gi 817130062 ref WP_046502581.1		
33.33 204 128 4 115 723 40 236 1e-26	116		
ORF_comp2951_c3_seq2	gi 1001837663 gb KXS32259.1	37.31	201 117
5 127 723 46 239 1e-26	116 ORF_comp2951_c3_seq2		
gi 800135199 dbj GAO34503.1	38.42 203 115 6 121 723 35 229		
2e-26 115 ORF_comp2951_c3_seq2	gi 918692141 ref WP_052561880.1		
33.99 203 118 4 130 735 40 227 3e-26	115		
ORF_comp2951_c3_seq2	gi 925230922 gb KOS75550.1	41.18	170 85
4 229 723 1 160 1e-25	111 ORF_comp2951_c3_seq2		
gi 522080484 ref WP_020591693.1	33.33 204 128 4 115 723 36		
232 2e-25 113 ORF_comp2951_c3_seq2	gi 497652379 ref WP_009966563.1		

[illegible]

33.67	199	106	7	160	723	83	266	9e-21	102										
ORF_comp2951_c3_seq2				gi 953489559 emb CEG46925.1				33.86		189	112								
5	160	723	188	364	9e-21		102	ORF_comp2951_c3_seq2											
gi 568057987 gb ETM54635.1						33.33			189	113	5	160	723	222	398				
1e-20		102	ORF_comp2951_c3_seq2		gi 570996289 gb ETP52930.1			33.33											
189	113	5	160	723	221	397	1e-20		102	ORF_comp2951_c3_seq2									
gi 675213985 ref XP_008915877.1						33.33			189	113	5	160	723	222					
398	1e-20		102	ORF_comp2951_c3_seq2		gi 570959958 gb ETP24937.1													
32.28		189	115	5	160	723	4	180	2e-20		99.0								
ORF_comp2951_c3_seq2				gi 970660098 gb KUF97679.1				33.33		189	113								
5	160	723	222	398	2e-20		101	ORF_comp2951_c3_seq2											
gi 800137352 dbj GAO33130.1						32.82		195	113	4	157	723	46	228					
2e-20		98.6	ORF_comp2951_c3_seq2		gi 570335908 gb ETO83856.1														
33.33		189	113	5	160	723	221	397	2e-20		101								
ORF_comp2951_c3_seq2				gi 544211776 ref XP_005536287.1				35.00		220									
111	9	124	735	137	340	2e-20		102	ORF_comp2951_c3_seq2										
gi 33187146 gb AAO13810.2 AF384111_1						34.56		217	123	8	115	750							
24	226	2e-20		102	ORF_comp2951_c3_seq2														
gi 737401579 ref WP_035382755.1						34.21		190	117	4	157	723	44						
226	2e-20		98.2	ORF_comp2951_c3_seq2		gi 397628930 gb EJK69115.1													
34.09		220	110	9	157	714	152	370	3e-20		100								
ORF_comp2951_c3_seq2				gi 675852512 ref XP_009011816.1				32.80		189									
114	4	160	723	114	290	3e-20		100	ORF_comp2951_c3_seq2										
gi 566032216 gb ETI55127.1						32.28		189	115	5	160	723	80	256					
3e-20		99.4	ORF_comp2951_c3_seq2		gi 570335911 gb ETO83859.1														
32.28		189	115	5	160	723	80	256	3e-20		99.4								
ORF_comp2951_c3_seq2				gi 570959955 gb ETP24934.1				32.80		189	114								
5	160	723	222	398	3e-20		100	ORF_comp2951_c3_seq2											
gi 568057990 gb ETM54638.1						32.28		189	115	5	160	723	80	256					
4e-20		99.0	ORF_comp2951_c3_seq2		gi 570996278 gb ETP52920.1														
32.28		189	115	5	160	723	80	256	4e-20		99.0								
ORF_comp2951_c3_seq2				gi 675213991 ref XP_008915880.1				32.28		189									
115	5	160	723	80	256	4e-20		99.0	ORF_comp2951_c3_seq2										
gi 511010013 gb EPB91244.1						32.09		215	119	9	115	732	111	307					
5e-20		101	ORF_comp2951_c3_seq2		gi 758368534 emb CEP09044.1			32.09											
215	119	9	115	732	112	308	5e-20		101	ORF_comp2951_c3_seq2									
gi 755598209 ref WP_042525907.1						32.75		171	108	3	229	735	1						
166	6e-20		95.9	ORF_comp2951_c3_seq2															
gi 695064415 ref XP_009420754.1						33.17		208	120	6	115	723	22						
215	6e-20		100	ORF_comp2951_c3_seq2		gi 312072379 ref XP_003139039.1													
29.67		209	128	5	115	723	29	224	8e-20		96.7								
ORF_comp2951_c3_seq2				gi 758352771 dbj GAN05107.1				31.92		213	122								
8	115	732	116	312	8e-20		100	ORF_comp2951_c3_seq2											
gi 499786267 ref WP_011467001.1						29.57		230	131	7	121	738	93						
315	8e-20		98.2	ORF_comp2951_c3_seq2		gi 970651949 gb KUF90697.1													
32.28		189	115	5	160	723	80	256	1e-19		97.8								
ORF_comp2951_c3_seq2				gi 373405317 gb AEY68568.1				33.03		218	123								
8	121	753	24	225	1e-19		99.8	ORF_comp2951_c3_seq2											
gi 223995009 ref XP_002287188.1						33.96		212	117	10	115	723	17						
214	2e-19		95.5	ORF_comp2951_c3_seq2															
gi 760447765 ref XP_011400951.1						32.38		210	120	6	160	765	60						
255	2e-19		98.2	ORF_comp2951_c3_seq2															
gi 831778908 ref XP_012755554.1						29.47		190	118	5	160	723	85						
260	2e-19		99.0	ORF_comp2951_c3_seq2															
gi 694415950 ref XP_009336117.1						32.11		218	125	8	121	753	36						
237	2e-19		99.0	ORF_comp2951_c3_seq2															
gi 242051649 ref XP_002454970.1						32.88		219	124	8	115	750	24						

[illegible]



gi 972777756 ref NP_001305597.1	33.33	201	114	7	160	750	4
188 3e-18	95.5 ORF_comp2951_c3_seq2						
gi 695071232 ref XP_009382679.1	31.66	199	120	5	160	750	40
224 3e-18	95.9 ORF_comp2951_c3_seq2	gi 325516458 gb ADZ24787.1					
33.85	192 107 7	160 723 40	215	3e-18	95.5		
ORF_comp2951_c3_seq2	gi 752845037 ref WP_041474048.1	30.36					224
133 5	121 726 31	253 3e-18	92.8	ORF_comp2951_c3_seq2			
gi 531558322 gb AGT57959.1	33.33	195 116 5	160 738 51	233			
3e-18	95.1 ORF_comp2951_c3_seq2	gi 848932098 ref XP_012828973.1					
33.83	201 113 7	160 750 40	224 3e-18	95.9			
ORF_comp2951_c3_seq2	gi 848932096 ref XP_012828972.1	33.83					201
113 7	160 750 40	224 3e-18	95.9	ORF_comp2951_c3_seq2			
gi 224965368 emb CAX56900.1	30.36	224 133 5	121 726 44	266			
3e-18	92.8 ORF_comp2951_c3_seq2	gi 727146285 emb CEG71517.1					
29.58	213 127 8	115 732 57	253 4e-18	95.1			
ORF_comp2951_c3_seq2	gi 568215457 ref NP_001275308.1	33.33					201
114 7	160 750 40	224 4e-18	95.5	ORF_comp2951_c3_seq2			
gi 857976683 emb CEO97478.1	29.73	185 115 3	154 702 37	208			
4e-18	94.7 ORF_comp2951_c3_seq2	gi 848932094 ref XP_012828971.1					
33.83	201 113 7	160 750 40	224 4e-18	95.9			
ORF_comp2951_c3_seq2	gi 1002820966 gb AMN87031.1	32.72					217 127
7 115 750 24	226 4e-18	95.5	ORF_comp2951_c3_seq2				
gi 926791675 ref XP_013904311.1	40.16	122 71 2	361 723 5				
125 4e-18	90.1 ORF_comp2951_c3_seq2	gi 672829171 gb KFH74060.1					
33.51	191 106 6	154 714 403 576 4e-18	95.9				
ORF_comp2951_c3_seq2	gi 670428756 ref XP_008655225.1	32.26					217
128 7	115 750 24	226 4e-18	95.1	ORF_comp2951_c3_seq2			
gi 297848618 ref XP_002892190.1	33.33	192 108 7	160 723 40				
215 4e-18	94.4 ORF_comp2951_c3_seq2	gi 66803020 ref XP_635353.1					
30.20	202 124 5	160 762 168 353 5e-18	95.5				
ORF_comp2951_c3_seq2	gi 293337247 ref NP_001168641.1	32.26					217
128 7	115 750 24	226 5e-18	95.1	ORF_comp2951_c3_seq2			
gi 302847090 ref XP_002955080.1	34.52	168 91 5	160 654 36				
187 5e-18	90.9 ORF_comp2951_c3_seq2						
gi 302782732 ref XP_002973139.1	31.71	205 116 7	160 762 40				
224 6e-18	94.7 ORF_comp2951_c3_seq2						
gi 302789570 ref XP_002976553.1	31.71	205 116 7	160 762 40				
224 6e-18	94.7 ORF_comp2951_c3_seq2	gi 674249977 gb KFK42742.1					
33.33	192 108 7	160 723 40	215 6e-18	94.4			
ORF_comp2951_c3_seq2	gi 312171048 emb CBX79307.1	29.91					214 127
5 121 696 17	229 6e-18	91.7	ORF_comp2951_c3_seq2				
gi 1002820968 gb AMN87032.1	31.51	219 127 8	115 750 24	226			
7e-18	94.7 ORF_comp2951_c3_seq2	gi 552916579 gb ESA01781.1					
30.41	217 113 6	115 723 17	209 8e-18	90.9			
ORF_comp2951_c3_seq2	gi 444891677 gb AGE13358.1	30.59					255 114
10 115 723 57	300 8e-18	92.8	ORF_comp2951_c3_seq2				
gi 478729310 emb CCP05801.1	29.91	214 127 5	121 696 31	243			
9e-18	91.3 ORF_comp2951_c3_seq2	gi 224004956 ref XP_002296129.1					
32.82	195 111 6	157 696 93	282 1e-17	92.4			
OR							

1e-17	94.0	ORF_comp2951_c3_seq2	gi 743835878 ref XP_010935937.1						
33.51	191	109	7	160	723	40	215	1e-17	93.6
ORF_comp2951_c3_seq2	gi 113594574 dbj BAF18448.1	31.82	198	121					
4	160	750	41	225	1e-17	90.9	ORF_comp2951_c3_seq2		
gi 15617416 dbj BAB64932.1	30.46	197	117	6	160	738	40	220	
1e-17	94.0	ORF_comp2951_c3_seq2	gi 552820766 ref XP_005845668.1						
30.29	175	107	4	154	675	47	207	1e-17	90.1
ORF_comp2951_c3_seq2	gi 515351240 ref WP_016863694.1	30.46	197						
122	5	154	738	37	220	1e-17	92.8	ORF_comp2951_c3_seq2	
gi 671683790 emb CDS12824.1	32.80	189	112	5	160	723	119	293	
1e-17	94.0	ORF_comp2951_c3_seq2	gi 916690104 ref WP_051297195.1						
27.85	237	133	7	115	732	40	269	1e-17	90.9
ORF_comp2951_c3_seq2	gi 502816264 ref WP_013051240.1	30.22	225						
129	8	121	723	38	258	1e-17	90.9	ORF_comp2951_c3_seq2	
gi 562746587 ref WP_023653709.1	29.91	214	127	5	121	696	31		
243	2e-17	90.9	ORF_comp2951_c3_seq2	gi 15219610 ref NP_171894.1					
33.33	192	108	7	160	723	39	214	2e-17	93.2
ORF_comp2951_c3_seq2	gi 743835874 ref XP_010935936.1	33.51	191						
109	7	160	723	40	215	2e-17	93.6	ORF_comp2951_c3_seq2	
gi 672825823 gb KFH70714.1	33.33	201	114	6	160	756	148	330	
2e-17	94.0	ORF_comp2951_c3_seq2	gi 505014688 ref WP_015201790.1						
29.83	181	113	4	154	693	37	204	2e-17	92.8
ORF_comp2951_c3_seq2	gi 727439853 ref XP_010501307.1	33.85	192						
107	7	160	723	40	215	2e-17	94.0	ORF_comp2951_c3_seq2	
gi 727439853 ref XP_010501307.1	33.71	178	98	7	202	723	259		
420	5e-14	83.6	ORF_comp2951_c3_seq2	gi 909144132 gb KNE68198.1					
30.70	215	129	6	115	738	136	337	2e-17	93.6
ORF_comp2951_c3_seq2	gi 550281002 ref WP_022607401.1	33.52	179						
100	3	157	693	65	224	2e-17	90.5	ORF_comp2951_c3_seq2	
gi 734553453 gb KHN80812.1	30.77	195	121	4	160	738	43	225	
2e-17	92.0	ORF_comp2951_c3_seq2	gi 520899875 ref WP_020322885.1						
28.70	223	135	6	121	723	32	252	2e-17	90.1
ORF_comp2951_c3_seq2	gi 971557981 ref XP_015166074.1	32.84	201						
115	7	160	750	17	201	3e-17	92.8	ORF_comp2951_c3_seq2	
gi 971557978 ref XP_015166073.1	32.84	201	115	7	160	750	40		
224	3e-17	92.8	ORF_comp2951_c3_seq2						
gi 122202937 sp Q2QKL5.1 PCS3_LOTJA	30.00	210	124	7	115	723			
22	215	3e-17	92.4	ORF_comp2951_c3_seq2					
gi 330794008 ref XP_003285073.1	29.35	201	127	5	154	750	74		
261	3e-17	92.8	ORF_comp2951_c3_seq2						
gi 224129604 ref XP_002320627.1	30.35	201	120	5	115	702	22		
207	3e-17	92.8	ORF_comp2951_c3_seq2	gi 965670943 dbj BAT85550.1					
31.10	209	123	6	115	723	23	216	4e-17	92.4
ORF_comp2951_c3_seq2	gi 444891679 gb AGE13359.1	32.31	195	112					
6	157	696	93	282	4e-17	90.9	ORF_comp2951_c3_seq2		
gi 19114464 ref NP_593552.1	28.27	191	123	5	154	723	73	250	
4e-17	91.7	ORF_comp2951_c3_seq2	gi 215983522 gb ACJ71777.1						
29.84	191	120	4	154	723	38	215	4e-17	89.0
ORF_comp2951_c3_seq2	gi 909141466 gb KNE66257.1	32.80	189	113					
4	160	723	87	262	4e-17	92.8	ORF_comp2951_c3_seq2		
gi 676493293 ref XP_009066193.1	29.67	182	114	4	160	702	34		
202	4e-								

gi 676430709 ref XP_009046015.1	29.63	189	120	4	160	723	46
222 6e-17	88.6 ORF_comp2951_c3_seq2						
gi 331218210 ref XP_003321783.1	31.31	198	117	5	154	723	178
364 6e-17	90.9 ORF_comp2951_c3_seq2	gi 902234844 gb KNA23610.1					
33.16	193 107 8	160 723 41	216	6e-17	92.0		
ORF_comp2951_c3_seq2	gi 50659119 gb AAT80341.1	31.16	199	114			
6 136 723 32 210	6e-17	91.7 ORF_comp2951_c3_seq2					
gi 388497750 gb AFK36941.1	30.00	210 124 7	115	723 22	215		
7e-17	91.7 ORF_comp2951_c3_seq2	gi 659119977 ref XP_008459945.1					
30.29	208 122 7	121 723 24	215	7e-17	91.7		
ORF_comp2951_c3_seq2	gi 573946675 ref XP_006655670.1	31.31	198				
122 4	160 750 42	226 8e-17	91.7 ORF_comp2951_c3_seq2				
gi 698514101 ref XP_009801957.1	33.33	201 114 7	160	750 40			
224 8e-17	91.7 ORF_comp2951_c3_seq2						
gi 567155800 ref XP_006418181.1	32.29	192 110 7	160	723 4			
179 8e-17	90.9 ORF_comp2951_c3_seq2						
gi 331218212 ref XP_003321784.1	32.02	203 111 5	154	723 51			
239 8e-17	90.9 ORF_comp2951_c3_seq2	gi 1001614567 gb KXS18316.1					
31.25	192 116 5	154 723 120	297	8e-17	89.4		
ORF_comp2951_c3_seq2	gi 698514104 ref XP_009801958.1	33.33	201				
114 7	160 750 40	224 8e-17	91.7 ORF_comp2951_c3_seq2				
gi 504984674 ref WP_015171776.1	29.89	184 115 4	154	702 39			
209 8e-17	90.5 ORF_comp2951_c3_seq2						
gi 156379434 ref XP_001631462.1	25.95	185 122 3	160	702 40			
213 9e-17	88.2 ORF_comp2951_c3_seq2	gi 768677235 gb AJU57239.1					
33.33	201 114 7	160 750 40	224	9e-17	91.3		
ORF_comp2951_c3_seq2	gi 985450620 ref XP_015386245.1	29.52	210				
125 7	115 723 22	215 9e-17	91.3 ORF_comp2951_c3_seq2				
gi 166798217 gb ABY89660.1	32.29	192 110 7	160	723 40	215		
9e-17	91.3 ORF_comp2951_c3_seq2	gi 502110105 ref XP_004493800.1					
31.69	183 108 4	160 702 39	206	9e-17	91.3		
ORF_comp2951_c3_seq2	gi 743869429 ref XP_010905810.1	30.92	207				
122 6	121 723 24	215 9e-17	91.3 ORF_comp2951_c3_seq2				
gi 734345461 gb KHN10759.1	30.62	209 124 6	115	723 21	214		
9e-17	90.9 ORF_comp2951_c3_seq2	gi 641841507 gb KDO60419.1					
29.52	210 125 7	115 723 20	213	1e-16	91.3		
ORF_comp2951_c3_seq2	gi 923123265 ref XP_013755101.1	32.18	202				
119 8	160 747 39	228 1e-16	90.9 ORF_comp2951_c3_seq2				
gi 567185112 ref XP_006403145.1	34.38	192 106 7	160	723 40			
215 1e-16	91.3 ORF_comp2951_c3_seq2	gi 21104516 dbj BAB93119.1					
32.81	192 109 7	160 723 40	215	1e-16	91.3		
ORF_comp2951_c3_seq2	gi 961088958 ref XP_014771294.1	32.46	191				
111 6	160 723 39	214 1e-16	90.9 ORF_comp2951_c3_seq2				
gi 674895889 emb CDY36967.1	30.37	191 117 4	115	678 22	199		
1e-16	90.5 ORF_comp2951_c3_seq2	gi 302843401 ref XP_002953242.1					
29.89	174 107 4	160 678 44	203	1e-16	87.8		
ORF_comp2951_c3_seq2	gi 567902606 ref XP_006443791.1	29.52	210				
125 7	115 723 22	215 1e-16	91.3 ORF_comp2951_c3_seq2				
gi 641841510 gb KDO60422.1	33.68	193 108 7	160	726 40	216		
1e-16	91.3 ORF_comp2951_c3_seq2	gi 950949688 ref XP_014495217.1					
30.62	209 124 6	115 723 23	216	1e-16	90.9		
ORF_comp2951_c3_seq2	gi 922429656 ref XP_013620881.1	33.33	192				
108 7	160 723 40	215 1e-16	90.9 ORF_comp2951_c3_seq2				
gi 674936482 emb CDX96939.1	33.33	192 108 7	160	723 40	215		
1e-16	90.9 ORF_comp2951_c3_seq2	gi 685383195 ref XP_009123930.1					
33.33	192 108 7	160 723 40	215	1e-16	90.9		
ORF_comp2951_c3_seq2	gi 675852508 ref XP_009011814.1	31.84	201				

122	5	154	750	80	267	1e-16	90.9	ORF_comp2951_c3_seq2											
gi 27448224 gb AAO13809.1 AF384110.1							32.81		192	109	7	160	723						
40	215	1e-16				90.9	ORF_comp2951_c3_seq2												
gi 922429654 ref XP_013620880.1							33.33		192	108	7	160	723	40					
215	1e-16					90.9	ORF_comp2951_c3_seq2												
gi 923864531 ref XP_013707638.1							33.33		192	108	7	160	723	40					
215	1e-16					90.9	ORF_comp2951_c3_seq2												
gi 567155803 ref XP_006418182.1							32.29		192	110	7	160	723	40					
215	1e-16					90.5	ORF_comp2951_c3_seq2												
gi 685951299 ref XP_009269521.1							31.22		189	115	5	160	723	57					
231	1e-16					90.1	ORF_comp2951_c3_seq2												
gi 568851639 ref XP_006479495.1							33.68		193	108	7	160	726	40					
216	1e-16					90.9	ORF_comp2951_c3_seq2												
gi 922563714 ref XP_013610130.1							30.57		193	114	5	115	678	24					
201	1e-16					90.9	ORF_comp2951_c3_seq2												
gi 694387379 ref XP_009369444.1							32.35		204	118	7	154	753	38					
225	1e-16					90.9	ORF_comp2951_c3_seq2		gi 313483731 gb ADR51694.1										
30.46		197	117	6	160	738	4	184	2e-16		90.5								
ORF_comp2951_c3_seq2					gi 763775184 gb KJB42307.1				32.81			192	109						
7	160	723	42	217	2e-16		88.6	ORF_comp2951_c3_seq2											
gi 567902610 ref XP_006443793.1							30.20		202	119	6	115	702	22					
207	2e-16					90.5	ORF_comp2951_c3_seq2												
gi 685326730 ref XP_009101739.1							31.58		190	114	6	160	723	40					
215	2e-16					90.5	ORF_comp2951_c3_seq2		gi 313483729 gb ADR51693.1										
30.26		195	120	5	160	738	4	184	2e-16		90.5								
ORF_comp2951_c3_seq2					gi 891585491 ref XP_013022696.1				28.27			191							
123	5	154	723	74	251	2e-16	90.1	ORF_comp2951_c3_seq2											
gi 1002277234 ref XP_015642955.1							31.82		198	121	4	160	750	41					
225	2e-16					90.5	ORF_comp2951_c3_seq2												
gi 4768281 gb AAD29446.1 AF085231.2							30.77		195	121	4	160	723						
45	232	2e-16				90.5	ORF_comp2951_c3_seq2												
gi 470107265 ref XP_004289969.1							30.73		218	128	8	121	753	24					

ORF_comp2951_c3_seq2	gi 313483717 gb ADR51687.1	30.26	195	120
5 160 738 4 184 2e-16	90.1 ORF_comp2951_c3_seq2			
gi 923847316 ref XP_013702842.1	31.77	192 111 7	160 723 40	
215 2e-16 88.2 ORF_comp2951_c3_seq2	gi 641841509 gb KDO60421.1			
33.68 193 108 7 160 726 40 216 2e-16	90.5			
ORF_comp2951_c3_seq2	gi 313483733 gb ADR51695.1	30.46	197	117
6 160 738 4 184 2e-16	90.1 ORF_comp2951_c3_seq2			
gi 47155943 gb AAT11885.1	30.26	195 120 5	160 738 4	184
2e-16 90.1 ORF_comp2951_c3_seq2	gi 641841508 gb KDO60420.1			
30.20 202 119 6 115 702 22 207 2e-16	90.1			
ORF_comp2951_c3_seq2	gi 697171098 ref XP_009594480.1	32.84		201
115 7 160 750 40 224 3e-16	90.1 ORF_comp2951_c3_seq2			
gi 313483747 gb ADR51702.1	30.26	195 120 5	160 738 4	184
3e-16 89.7 ORF_comp2951_c3_seq2	gi 313483737 gb ADR51697.1			
30.26 195 120 5 160 738 4	184 3e-16	89.7		
ORF_comp2951_c3_seq2	gi 313483751 gb ADR51704.1	30.46	197	117
6 160 738 4 184 3e-16	89.7 ORF_comp2951_c3_seq2			
gi 313483713 gb ADR51685.1	30.26	195 120 5	160 738 4	184
3e-16 89.7 ORF_comp2951_c3_seq2	gi 313483711 gb ADR51684.1			
30.26 195 120 5 160 738 4	184 3e-16	89.7		
ORF_comp2951_c3_seq2	gi 763775185 gb KJB42308.1	32.81	192	109
7 160 723 42 217 3e-16	89.0 ORF_comp2951_c3_seq2			
gi 356505041 ref XP_003521301.1	30.62	209 124 6	115 723 21	
214 3e-16 90.1 ORF_comp2951_c3_seq2	gi 313483725 gb ADR51691.1			
30.46 197 117 6 160 738 4	184 3e-16	89.7		
ORF_comp2951_c3_seq2	gi 313483715 gb ADR51686.1	30.46	197	117
6 160 738 4 184 3e-16	89.7 ORF_comp2951_c3_seq2			
gi 657981621 ref XP_008382837.1	32.84	204 117 7	154 753 38	
225 3e-16 89.7 ORF_comp2951_c3_seq2	gi 674877470 emb CDY54482.1			
31.77 192 111 7 160 723 40 215 3e-16	89.7			
ORF_comp2951_c3_seq2	gi 313483721 gb ADR51689.1	30.46	197	117
6 160 738 4 184 3e-16	89.7 ORF_comp2951_c3_seq2			
gi 207100019 emb CAK24968.2	31.77	192 111 7	160 723 40	215
3e-16 89.7 ORF_comp2951_c3_seq2	gi 13928024 emb CAC37692.1			
31.77 192 111 7 160 723 40 215 3e-16	89.7			
ORF_comp2951_c3_seq2	gi 676476523 ref XP_009060820.1	28.02		182
117 4 160 702 24 192 3e-16	85.9 ORF_comp2951_c3_seq2			
gi 923885023 ref XP_013714028.1	29.53	193 116 5	115 678 22	
199 3e-16 89.7 ORF_comp2951_c3_seq2	gi 475597402 gb EMT23611.1			
34.39 189 104 7 160 714 40 212 3e-16	89.7			
ORF_comp2951_c3_seq2	gi 170594708 ref XP_001902100.1	31.43		210
110 7 160 738 44 236 3e-16	89.0 ORF_comp2951_c3_seq2			
gi 242051651 ref XP_002454971.1	31.80	217 128 8	121 750 26	
229 4e-16 89.4 ORF_comp2951_c3_seq2	gi 870854054 gb KMT05873.1			
31.00 200 118 7 160 723 41 232 4e-16	89.7			
ORF_comp2951_c3_seq2	gi 313483704 gb ADR51681.1	30.46	197	117
6 160 738 4 184 4e-16	89.4 ORF_comp2951_c3_seq2			
gi 902234843 gb KNA23609.1	32.04	206 116 9	160 762 41	227
4e-16 89.7 ORF_comp2951_c3_seq2				
gi 147719992 sp Q2TE74.2 PCS2_LOTJA	31.00	200 118 6	136 723	
32 215 4e-16 89.4 ORF_comp2951_c3_seq2				
gi 657981623 ref XP_008382838.1	32.84	204 117 7	154 753 29	
216 4e-16 89.4 ORF_comp2951_c3_seq2	gi 902234842 gb KNA23608.1			
32.64 193 108 8 160 723 41 216 4e-16	89.4			
ORF_comp2951_c3_seq2	gi 922533447 ref XP_013598130.1	31.77		192
111 7 160 723 96 271 4e-16	89.7 ORF_comp2951_c3_seq2			
gi 870854055 gb KMT05874.1	31.77	192 111 7	160 723 41	216

[illegible]

gi 470240016 ref XP_004351977.1	28.80	191	119	6	160	723	66
242 9e-16	88.2 ORF_comp2951_c3_seq2	gi 53760453 gb AAU93349.1					
32.29	192 110 7 160 723 40 215 9e-16	88.6					
ORF_comp2951_c3_seq2	gi 915844541 ref WP_050904055.1	29.03	217				
132 7 133 723 36 250 9e-16	85.5 ORF_comp2951_c3_seq2						
gi 672183478 ref XP_008812019.1	31.86	204 117 6 160 762 40					
224 9e-16	88.6 ORF_comp2951_c3_seq2	gi 966806441 gb ALT55649.1					
32.81	192 109 7 160 723 40 215 9e-16	88.6					
ORF_comp2951_c3_seq2	gi 449445620 ref XP_004140570.1	29.95	207				
121 7 127 723 25 215 9e-16	88.2 ORF_comp2951_c3_seq2						
gi 363413051 gb AEW23125.1	31.77	192 111 7 160 723 40 215					
9e-16	88.2 ORF_comp2951_c3_seq2	gi 674920164 emb CDY13128.1					
32.28	189 114 4 160 723 40 215 9e-16	88.2					
ORF_comp2951_c3_seq2	gi 21104518 dbj BAB93120.1	32.29	192 110				
7 160 723 40 215 1e-15	88.2 ORF_comp2951_c3_seq2						
gi 823127933 ref XP_012440917.1	28.86	201 123 5 115 702 22					
207 1e-15	88.2 ORF_comp2951_c3_seq2						
gi 743862902 ref XP_010943664.1	33.17	205 113 7 160 762 40					
224 1e-15	88.2 ORF_comp2951_c3_seq2	gi 763743549 gb KJB11048.1					
28.86	201 123 5 115 702 22 207 1e-15	88.2					
ORF_comp2951_c3_seq2	gi 46949222 gb AAT07467.1	31.77	192 111				
7 160 723 40 215 1e-15	88.2 ORF_comp2951_c3_seq2						
gi 926792009 ref XP_013904478.1	30.16	189 116 5 160 720 24					
198 1e-15	87.8 ORF_comp2951_c3_seq2	gi 947119173 gb KRH67422.1					
30.89	191 114 5 160 723 4 179 1e-15	87.8					
ORF_comp2951_c3_seq2	gi 693266154 gb AIS24729.1	29.85	201 118				
6 154 723 69 257 1e-15	87.8 ORF_comp2951_c3_seq2						
gi 645260011 ref XP_008235633.1	31.86	204 119 7 154 753 38					
225 1e-15	88.2 ORF_comp2951_c3_seq2	gi 388518807 gb AFK47465.1					
31.09	193 115 5 154 723 40 217 1e-15	88.2					
ORF_comp2951_c3_seq2	gi 545355774 ref XP_005643662.1	30.16	189				
112 6 160 720 17 187 1e-15	84.3 ORF_comp2951_c3_seq2						
gi 122208907 sp Q2TSC7.1 PCS1_LOTJA	31.09	193 115 5 154 723					
38 215 1e-15	88.2 ORF_comp2951_c3_seq2						
gi 384491545 gb EIE82741.1	30.99	213 98 8 115 732 53 223					
1e-15	87.4 ORF_comp2951_c3_seq2	gi 590715531 ref XP_007050222.1					
30.20	202 119 6 115 702 22 207 1e-15	87.8					
ORF_comp2951_c3_seq2	gi 922548834 ref XP_013602398.1	31.77	192				
111 7 160 723 40 215 1e-15	87.8 ORF_comp2951_c3_seq2						
gi 313483719 gb ADR51688.1	30.46	197 117 6 160 738 4 184					
1e-15	87.8 ORF_comp2951_c3_seq2	gi 220683835 gb ACL80669.1					
32.29	192 110 7 160 723 40 215 1e-15	87.8					
ORF_comp2951_c3_seq2	gi 922346013 ref XP_013449921.1	31.15	183				
109 4 160 702 41 208 1e-15	87.8 ORF_comp2951_c3_seq2						
gi 255964729 gb ACU44656.1	30.73	192 113 7 160 723 40 215					
1e-15	87.8						

Horizontal gene transfer of phytochelatin synthases from bacteria to extremophilic green algae; Microbial Ecology; Olsson S\*, Penacho V, Puente-Sánchez F, Díaz S, Aguilera A; \* corresponding author, Department of Agricultural Sciences, P.O. Box 27, 00014 University of Helsinki, Finland, sanna.olsson@helsinki.fi

**Online resource 5** Introns in the predicted PCS genes from *C. acidophila* and *D. acidophila*. There are three spliceosomal introns in CaPCS2, four in CaPCS1 and three in DaPCS1. The introns are either the most common GT-AG splice site or variations of it.

Gene	Intron 5' end	Intron 3' end
CaPCS2	GAG   GTGAGG	TAG   G
	CAG   GTGTGT	TAN   G
	CAA   GTAAGG	TAG   G
CaPCS1	AAG   GTAAGC	CAG   G
	CAG   GTACTT	CAG   C
	CAG   GTGATG	CAG   A
	CAG   GTGAGA	TAG   G
DaPCS1	CTG   GTGGGT	GAG   G
	CAG   GCGAGC	CAG   A
	CAA   GTGCGG	CAG   G



Horizontal gene transfer of phytochelatin synthases from bacteria to extremophilic green algae; Microbial Ecology;  
Olsson S\*, Penacho V, Puente-Sánchez F, Díaz S, Aguilera A;  
\* corresponding author, Department of Agricultural Sciences, P.O. Box 27,  
00014 University of Helsinki, Finland, sanna.olsson@helsinki.fi.

Online Resource 6 Isoforms of component comp17619\_c0 coding for the CaPCS1 gene in *Chlamydomonas acidophila*

>comp17619\_c0\_seq1

GCTCAATGTTAGAACCGCAAAATTTTGCTGTGGCTACCTTAAAAGTTTTTCAGGTCTAAAATCCAGTAAGTTCA  
TGTCCA  
AGCTGTGCATACCAAACCTGCTTATGTATCACAGTCTTATCCCATGTCCAGGTGCTATCACACCCTCCAATTTT  
GAGAAA  
CGACACGCAACAATATCTTTTTTCGGGCACTCCTTCTTCCTTGAAGGGAATTGAATACAACAGCGCCACATTGTC  
TCCAAT  
GAATGAAACAGTTCCTGAGATGACATGCAGCACTCAAAAATCAAGCTGTACCTCTAGGGTTGAGCATGTTTGTG  
GAGAGC  
CACCTCTTCAAAAAATCCTATCTCAATACACCTGTATAATATCTCTGGTGCCGGGTCTGTGCAACCTTCTCAA  
TCAACA  
GCCACTCCTATTGTGGGGGTGACAGCTCTTCAAACAGGCTCTGGGTTTAGAAGGACATTTTATAAGCGTCATCT  
ACCAAG  
CCCACCAGCAATTGCTTTTTCTCACCTCAAGGGCGTCAAATATTTCAAGAGGCGCTAATGGGCGGCACCCTAG  
GAGGCT  
TTTTCAAGTTGATGGAGCAATTCAGCATGCAAGACGAACCAGCCTTTTGTGGTCTTACAAGCTTGACTATGGTG  
CTGAAT  
GCACTCTCCATCGACCCGCGGCGAACATGGAAGGTGCATGGCGTTGGTTCATGAGGCCATGCTAGACTGCTGC  
AGGTGA  
TGAGTCACCAAATCAAACGCTGTGATTCTCAGCTCCTGATGCTCACTTTTCAACTACATCTTAATCCCATCCC  
CATCCA  
TCGCAACAGCTTATGACCCTCATATCAACTAGAAGTGATATGGAACTTTGCTTTCTCTCTCACAGACCATTGG  
AATCTG  
TTAAGGAGGATGGCATAACCCTTGCTCAGGCTGCATGTCTTGCCCGCTGCAACGGGGCCCGCGTTGACATGATG  
CGACAT  
GGCACCTTCTCAGAGGCTGAGTTCAGGCAGCAAGTGATGGATGTGTGTAGGAGTGGGGAGGAGCATCTGGTGGT  
TAGCTA  
CGCCCGACGTGCTTTTAGTCAGACTGGGGATGGTCATTTTAGTCCTGTTGGAGGATACCATGAGGGCCAGGACT  
TGGTAC  
TAATCTTGGATGTGGCCCGCTTCAAGTATCCTCCACACTGGGTACCTCTACCTCTGCTCTATGAGGCCATGTCA  
CATGTA  
GACCCCAACCTGGCCTCCACGTGGCTACATGCTCCTGTCTGCTCACCCGCTGCTGGACTCCGCCATGTTCTC  
TCTAGA  
CATCCGTCAAGAGGGATGGAAGCTGGCAAGAAGCTATGTCACCATGATGCCTGTAGTCTTGGAGCAGTACATGT  
CCTTGA  
GAGCTCAGAGGCTGCAGCAGGAGCAGCAGATGTTGGCAGCCAAGGAAGACAAGCTTGTGGTTATGGAGCATGCT  
GTAAGG  
CATTTAGTCCTAGGTATTCCAAAACAAGGCATTATATCCTTCATTGCCATTCGGGAGCTGGCAAGAGCGCCTGG  
AGAGAA  
GTGTGTGCCTCAGAAGTCTAGAGAGACTCTATTGTTGGAGCTGAGGGCCATGGTTTGTATCAGCTGGTGCTCAA  
GGCAAT  
CAAAGGGGATGGACCATCTGCCTCGGCAGCAGCTGATAGAGAAGCTGGTATGCGGAGCCCTGAGGATACAGAGG  
ACATGG  
GCGGATCTCTGTGTTCTTCCTCGGCCTACATGGCAGATCGTTGTACACTTCTCATCATGCTGCAGCCCCCTCA  
GCATGG  
CCAGCCTGGTCATCTCCAGAGCTCAACACTCAGTGGCAAGCTCTTTTAGATCTAACGAATTTCAAGTGTGTTGC  
TACTGA  
GGTGGCTTACATGCGCCAGCAGTTTGTACATGTGGATGAAGTGATGCAGGGAGACTCAGTGGACGCTACTTGTG

GCCAAC  
CTGCATGCAAATCCAGACTCGACTGTGCAGTTCTGGGACATTGACTTTGAATGACTGCTTCATGCTGAGAACAA  
CTGTGA  
TTCTGTTGGCATGACTGAGCCAATCAATATCTCTTTTGATAAACCAGACTGAATATACATTATGGGTTTGAGCA  
AATGCC  
CATAAGAGACTTTGTGCACGTCCAACATCCTTAGTCACTGTGACTAATGAGACTTTAAACCGCTCTCCTTCGCA  
CATAG

>comp17619\_c0\_seq2

GCTCAATGTTAGAACCGCAAAATTTTGCTGTGGCTACCTTAAAAGTTTTTCAGGTCTAAAATCCAGTGCTATCA  
CACCCT  
CCAATTTTGAGAAACGACACGCAACAATATCTTTTTCGGGCACTCCTTCTTCCTTGAAGGGAATTGAATACAAC  
AGCGCC  
ACATTGTCTCCAATGAATGAAACAGTTCCTGAGATGACATGCAGCACTCAAAAATCAAGCTGTACCTCTAGGGT  
TGAGCA  
TGTTTGTGGAGAGCCACCCTCTTCAAAAAATCCTATCTCAATACACCTGTATAATATCTCTGGTGCCGGGTCTG  
TGCAAC  
CTTCTCAATCAACAGCCACTCCTATTGTGGGGGTGACAGCTCTTCAAACAGGCTCTGGGTTTAGAAGGACATTT  
TATAAG  
CGTCATCTACCAAGCCCACCAGCAATTGCTTTTTCTCACCTCAAGGGCGTCAAATATTTCAAGAGGCGCTAAT  
GGGCGG  
CACCCTAGGAGGCTTTTTCAAGTTGATGGAGCAATTCAGCATGCAAGACGAACCAGCCTTTTGTGGTCTTACAA  
GCTTGA  
CTATGGTGCTGAATGCACTCTCCATCGACCCGCGGCGAACATGGAAGGTGCATGGCGTTGGTTCATGAGGCCA  
TGCTAG  
ACTGCTGCAGGTGATGAGTCACCAAATCAAACGCTGTGATTCTCAGCTCCTGATGCTCACTTTTCAACTACAT  
CTTAAT  
CCCATCCCCATCCATCGCAACAGCTTATGACCCTCATATCAACTAGAAAGTGATATGGAACTTTGCTTTCTCTC  
TCACAG  
ACCATTGGAATCTGTTAAGGAGGATGGCATAACCCTTGCTCAGGCTGCATGTCTTGCCCGCTGCAACGGGGCCC  
GCGTTG  
ACATGATGCGACATGGCACCTTCTCAGAGGCTGAGTTCAGGCAGCAAGTGATGGATGTGTGTAGGAGTGGGGAG  
GAGCAT  
CTGGTGGTTAGCTACGCCCCGACGTGCTTTTAGTCAGACTGGGGATGGTCATTTTAGTCCTGTTGGAGGATACCA  
TGAGGG  
CCAGGACTTGGTACTAATCTTGATGTGGCCCGCTTCAAGTATCCTCCACACTGGGTACCTCTACCTCTGCTCT  
ATGAGG  
CCATGTCACATGTAGACCCCACTGGCCTCCACGTGGCTACATGCTCCTGTCTGCTCACCCGCTGCTGGAC  
TCCGCC  
ATGTTCTCTCTAGACATCCGTCAAGAGGGATGGAAGCTGGCAAGAAGCTATGTCACCATGATGCCTGTAGTCTT  
GGAGCA  
GTACATGTCCTTGAGAGCTCAGAGGCTGCAGCAGGAGCAGCAGATGTTGGCAGCCAAGGAAGACAAGCTTGTGG  
TTATGG  
AGCATGCTGTAAGGCATTTAGTCCTAGGTATTCCAAAACAAGGCATTATATCCTTCATTGCCATTCGGGAGCTG  
GCAAGA  
GCGCCTGGAGAGAAGTGTGTGCCTCAGAAGTCTAGAGAGACTCTATTGTTGGAGCTGAGGGCCATGGTTTGTAT  
CAGCTG  
GTGCTCAAGGCAATCAAAGGGGATGGACCATCTGCCTCGGCAGCAGCTGATAGAGAAGCTGGTATGCGGAGCCC  
TGAGGA  
TACAGAGGACATGGGCGGATCTCTGTGTTCTTCCTCGGCCTACATGGCAGATCGTTGTACACTTCTCATCATGC  
TGCAGC  
CCCCCTCAGCATGGCCAGCCTGGTCATCTCCAGAGCTCAACACTCAGTGGCAAGCTCTTTTAGATCTAACGAAT  
TTCAGT  
GTTGTTGCTACTGAGGTGGCTTACATGCGCCAGCAGTTTGTACATGTGGATGAAGTGATGCAGGGAGACTCAGT  
GGACGC  
TACTTGTGGCCAACCTGCATGCAAATCCAGACTCGACTGTGCAGTTCTGGGACATTGACTTTGAATGACTGCTT  
CATGCT

GAGAACAAC TGTGATTCTGTTGGCATGACTGAGCCAATCAATATCTCTTTTGATAAAC CAGACTGAATATACAT  
TATGGG  
TTTGAGCAAATGCCCATAAGAGACTTTGTGCACGTCCAACATCCTTAGTCACTGTGACTAATGAGACTTTAAAC  
CGCTCT CCTTCGCACATAG

>comp17619\_c0\_seq3  
GCTCAATGTTAGAACCGCAAAATTTTGTCTGTGGCTACCTTAAAAGTTTTTCAGGTCTAAAATCCAGTAAGTTCA  
TGTCCA  
AGCTGTGCATACCAAAC TGTATGTATCACAGTCTTATCCCCATGTCCAGGTGCTATCACACCCTCCAATTTT  
GAGAAA  
CGACACGCAACAATATCTTTTTTCGGGCACTCCTTCTTCCTTGAAGGGAATTGAATACAACAGCGCCACATTGTC  
TCCAAT  
GAATGAAACAGTTCCTGAGATGACATGCAGCACTCAAAAATCAAGCTGTACCTCTAGGGTTGAGCATGTTTGTG  
GAGAGC  
CACCTCTTCAAAAAATCCTATCTCAATACACCTGTATAATATCTCTGGTGCCGGGTCTGTGCAACCTTCTCAA  
TCAACA  
GCCACTCCTATTGTGGGGGTGACAGCTCTTCAAACAGGCTCTGGGTTTAGAAGGACATTTTATAAGCGTCATCT  
ACCAAG  
CCCACCAGCAATTGCTTTTTCTCACCTCAAGGGCGTCAAATATTTCAAGAGGCGCTAATGGGCGGCACCCTAG  
GAGGCT  
TTTTCAAGTTGATGGAGCAATTCAGCATGCAAGACGAACCAGCCTTTTGTGGTCTTACAAGCTTGACTATGGTG  
CTGAAT  
GCACTCTCCATCGACCCGCGGCGAACATGGAAGGTGCATGGCGTTGGTTCCATGAGGCCATGCTAGACTGCTGC  
AGACCA  
TTGGAATCTGTCAAGGAGGATGGCATAACCCTTGCTCAGGCTGCATGTCTTGCCCGCTGCAACGGGGCCCGCGT  
TGACAT  
GATGCGACATGGCACCTTCTCAGAGGCTGAGTTCAGGCAGCAAGTGATGGATGTGTGTAGGAGTGGGGAGGAGC  
ATCTGG  
TGGTTAGCTACGCCC GACGTGCTTTTAGTCAGACTGGGGATGGTCATTTTAGTCCTGTTGGAGGATACCATGAG  
GGCCAG  
GACTTGGTACTAATCTTGATGTGGCCCGCTTCAAGTATCCTCCACACTGGGTACCTCTACCTCTGCTCTATGA  
GGCCAT  
GTCACATGTAGACCCCACTGGCCTCCACGTGGCTACATGCTCCTGTCTGCTCACCCGCTGCTGGACTCCG  
CCATGT  
TCTCTCTAGACATCCGTCAAGAGGGATGGAAGCTGGCAAGAAGCTATGTCACCATGATGCCTGTAGTCTTGGAG  
CAGTAC  
ATGTCCTTGAGAGCTCAGAGGCTGCAGCAGGAGCAGCAGATGTTGGCAGCCAAGGAAGACAAGCTTGTGGTTAT  
GGAGCA  
TGCTGTAAGGCATTTAGTCCTAGGTATTCCAAAACAAGGCATTATATCCTTCATTGCCATTCGGGAGCTGGCAA  
GAGCGC  
CTGGAGAGAAGTGTGTGCCTCAGAAGTCTAGAGAGACTCTATTGTTGGAGCTGAGGGCCATGGTTTGTATCAGC  
TGGTGC  
TCAAGGCAATCAAAGGGGATGGACCATCTGCCTCGGCAGCAGCTGATAGAGAAGCTGGTATGCGGAGCCCTGAG  
GATACA  
GAGGACATGGGCGGATCTCTGTGTTCTTCCTCGGCCTACATGGCAGATCGTTGTACACTTCTCATCATGCTGCA  
GCCCCC  
CTCAGCATGGCCAGCCTGGTCATCTCCAGAGCTCAACACTCAGTGGCAAGCTCTTTTAGATCTAACGAATTTCA  
GTGTTG  
TTGCTACTGAGGTGGCTTACATGCGCCAGCAGTTTGTACATGTGGATGAAGTGATGCAGGGAGACTCAGTGGAC  
GCTACT  
TGTGGCCAACCTGCATGCAAATCCAGACTCGACTGTGCAGTTCTGGGACATTGACTTTGAATGACTGCTTCATG  
CTGAGA  
ACAACTGTGATTCTGTTGGCATGACTGAGCCAATCAATATCTCTTTTGATAAAC CAGACTGAATATACATTATG  
GGTTTG  
AGCAAATGCCCATAAGAGACTTTGTGCACGTCCAACATCCTTAGTCACTGTGACTAATGAGACTTTAAACCGCT  
CTCCTT CGCACATAG

>comp17619\_c0\_seq4

CCGCTCTCCTTCGCACATAGGGACACGCCTCCCAGGAAGGTTTGCACACATCATATCCTCTAACAACATATTTT  
CTTTCA  
TTCTTATTTACATGACAAATATAACACATAGTCATCAAAAATTGAGGGATCGCAATCAGAGGCCTTCACTTATA  
TTTTCA  
TTTTATTATTACAGATCTCAACACTTGGTGCGCATTTTGGTCTCAGTGTTATTTTAATCACATTACTCCCAA  
ACATAT  
TTCTTAAACATGGCCTTCAAGAACCTTGCCAGACCTAGGACAATGACACACATTTTACAACGCCCTAAGCAGTC  
GTTGTA  
AACAACTACTCTTATTACTCTTACTCTCAACGGGCTCAAGCCATCCTGCAGTAGGCTAATAAAAGATAATCCATC  
TAACAC  
ATTGAGAATATTACACCATTTTCATTGCAGAGACGGACCTGCCTAGACTTCTGCAAGCTGGTGATGTCCAGAAAG  
AGCTTT  
TATCCTTCCTGATGAAGGCGGCACAAAACAACCGCAACAGGTTTACGAAGAAGTAGTTAGCCCTGCTGCTTATG  
CAGCCT  
GTTACTCATTTCTGCGATCTTTTTTCATATCTTCGGCTTCTTCCTGCAACTTGGCCACACACCACTATCTCCA  
TGCGCC  
ACATCCAATTATTCTCAGCTCGCCCAGGAGTATTCATTCTTGCAAGTATTATCTAATCTCATCACATCCTGCATC  
ATGACA  
ATGCATGTGGCTGCCACGCTGTTTCATGCAAGCTCGCAGCAGGTCCCAAGCAATATCATCTCCAGTAGTTTGTAG  
GTATTG  
ACACAGGTACTCCTTGCTTCCTTGGAAGCGCTGTCTTTAAACCAGCCCACCGATGTCTCATTGTCGTGGGTGC  
CAGGAT  
ACACAAAACAATTAGCATAGTGGTTGTGAGGGAGGTGGACATTGCCTGGGCCACCTCCCCATGCAACTGGAGC  
ACCACC  
ATGCCAGGAGCACCAATACTCTCCCTCAACTCCACCACGTCAGTGGTGATGACCCCTAGATCCTCCGCCAAGAT  
CGGCAC  
TGCCCCTAGCTTGGCCTTGATAGCTGTGAACAGCTCCAATCCGGGCCCTTTACACCACGACCCCCCATGGCAG  
TCTCAG  
CTTTACCATCAACAGCCCAATAGCCTGCAAAGCCCCGGAAGTGATCAATCCTTGTTTGGTCATAAAGCTGCATA  
GCTCGC  
CCCATGCGTTGACACCACCAGTCGTAGTTCTGCGCCTTGTGAGCCGACCACTTGTAAGCGGACTCCCCACAA  
CTGACC  
AGTCTCTGAGAATGCATCAGGTGGCACCCCCACTGACATTAGCTGGCAGGCCTGTAGGTCCTAGCTCAAAGAGGT  
CTTGAT  
TAGACCACACATCTGCTGAGTGCCACCCACATAAATGGGCATGTCTCCAATGATGTCCACACCTTTTGAGTTT  
GCATAC  
GCTCTGATAGCCTTCCACTGACGATCAAAGAAGAACTGTATGGCAATAAATTCCTCAATCGCCTGCTTATGTTT  
CTCCTG  
GAATGTTTTTCAGAGCATCCGGGTCCCTGAAGCGCAGCGACTCTGGCCAATCCCACCAGGCCATCGTAGATAGTT  
CTGAAG  
GTTCTTGCTACATCAAACACAGCGCTTTCCTCAATCCATGGATTCAATTTCTTAAATGCTATCATCTCCTCCC  
TGAGAT  
CCTTAAATTTGATGTCTTCTAGGAGCCTTTCAGCTGCAATGCGCAAGGCAGGCGTTTTAGCTGCCATGACAGCA  
GGGTAG  
TCAATGTTGCCAACTGAGACTTTAGCAGGGGCGTCCTTGGCATGTATCAAGCCCTCTTGAATCAACTCATCAAT  
AGAGAT  
GATGAGTGGATTGCCACAATTTGCGTCAGTTCCACTGTAAGGAGAGTAAAACATGGGGTCTGGAGGCCCAAGTG  
GTAGCA  
TCTGCCAGCAAGTCATCCCTGCTGTAACAAGCCAGTCTACAAACAGTCTTGCCTCTCCCCCAACTCGCCACAC  
CATAAG  
GTCCTGGTAGAGAAGTTGGATGCAAAATAATCCCAGCCCTCCGTTTAATTGCATTTTTTGGGTTGTAAACTGTC  
ATATGC  
TTGCGGTAGTGTATCCCCTATTGAGGCTTTCCTATAAACTGATGACATGTTTAATGCAGGTGCCTGGGCCACG  
GCAGTG  
CCCAAGTCTGTACTTTTCATCTTTGATGTTGAAATTTTGGCATCACGATGCATGGAGACAGCACACGATTTCTA  
CGTATA

TCATGCGCTAACAAACACGATCTAGCTTTGATCTGTTCTGATTTGATGAAGAGTTTCGACATCAGAATATCTAG  
TGAACA  
AACACAATATCTATGTTCTGATCACACAGAAATTATCAATGTATATTATTATCTTTAGTCTATTCAAGCATCGT  
TAAG

>comp17619\_c0\_seq5

GCTCAATGTTAGAACCGCAAAATTTTGCTGTGGCTACCTTAAAAGTTTTTCAGGTCTAAAATCCAGTGCTATCA  
CACCTT  
CCAATTTTGAGAAACGACACGCAACAATATCTTTTTTCGGGCACTCCTTCTTCCTTGAAGGGAATTGAATACAAC  
AGCGCC  
ACATTGTCTCCAATGAATGAAACAGTTCCTGAGATGACATGCAGCACTCAAAAATCAAGCTGTACCTCTAGGGT  
TGAGCA  
TGTTTGTGGAGAGCCACCCTCTTCAAAAAATCCTATCTCAATACACCTGTATAATATCTCTGGTGCCGGGTCTG  
TGCAAC  
CTTCTCAATCAACAGCCACTCCTATTGTGGGGGTGACAGCTCTTCAAACAGGCTCTGGGTTTAGAAGGACATTT  
TATAAG  
CGTCATCTACCAAGCCCACCAGCAATTGCTTTTTCTCACCTCAAGGGCGTCAAATATTTCAAGAGGCGCTAAT  
GGGCGG  
CACCTAGGAGGCTTTTTCAAGTTGATGGAGCAATTCAGCATGCAAGACGAACCAGCCTTTTGTGGTCTTACAA  
GCTTGA  
CTATGGTGCTGAATGCACTCTCCATCGACCCGCGGCGAACATGGAAGGTGCATGGCGTTGGTTCCATGAGGCCA  
TGCTAG  
ACTGCTGCAGACCATTGGAATCTGTCAAGGAGGATGGCATAACCCTTGCTCAGGCTGCATGTCTTGCCCGCTGC  
AACGGG  
GCCCgcgTTGACATGATGCGACATGGCACCTTCTCAGAGGCTGAGTTCAGGCAGCAAGTGATGGATGTGTGTAG  
GAGTGG  
GGAGGAGCATCTGGTGGTTAGCTACGCCCGACGTGCTTTTAGTCAGACTGGGGATGGTCATTTTAGTCCTGTTG  
GAGGAT  
ACCATGAGGGCCAGGACTTGGTACTAATCTTGATGTGGCCCGCTTCAAGTATCCTCCACACTGGGTACCTCTA  
CCTCTG  
CTCTATGAGGCCATGTCACATGTAGACCCCACTGGCCTCCACGTGGCTACATGCTCCTGTCTGCTCACCC  
GCTGCT  
GGACTCCGCCATGTTCTCTCTAGACATCCGTCAAGAGGGATGGAAGCTGGCAAGAAGCTATGTCACCATGATGC  
CTGTAG  
TCTTGGAGCAGTACATGTCCTTGAGAGCTCAGAGGCTGCAGCAGGAGCAGCAGATGTTGGCAGCCAAGGAAGAC  
AAGCTT  
GTGGTTATGGAGCATGCTGTAAGGCATTTAGTCCTAGGTATTCCAAAACAAGGCATTATATCCTTCATTGCCAT  
TCGGGA  
GCTGGCAAGAGCGCCTGGAGAGAAGTGTGTGCCTCAGAAGTCTAGAGAGACTCTATTGTTGGAGCTGAGGGCCA  
TGGTTT  
GTATCAGCTGGTGCTCAAGGCAATCAAAGGGGATGGACCATCTGCCTCGGCAGCAGCTGATAGAGAAGCTGGTA  
TGCGGA  
GCCCTGAGGATACAGAGGACATGGGCGGATCTCTGTGTTCTTCCTCGGCCTACATGGCAGATCGTTGTACACTT  
CTCATC  
ATGCTGCAGCCCCCTCAGCATGGCCAGCCTGGTCATCTCCAGAGCTCAACACTCAGTGGCAAGCTCTTTTAGA  
TCTAAC  
GAATTTcAGTGTTGTTGCTACTGAGGTGGCTTACATGCGCCAGCAGTTTGTACATGTGGATGAAGTGATGCAGG  
GAGACT  
CAGTGGACGCTACTTGTGGCCAACCTGCATGCAAATCCAGACTCGACTGTGCAGTTCTGGGACATTGACTTTGA  
ATGACT  
GCTTCATGCTGAGAACAACCTGTGATTCTGTTGGCATGACTGAGCCAATCAATATCTCTTTTGATAAACCACT  
GAATAT  
ACATTATGGGTTTGAGCAAATGCCATAAGAGACTTTGTGCACGTCCAACATCCTTAGTCACTGTGACTAATGA  
GACTTT AAACCGCTCTCCTTCGCACATAG

>comp17619\_c0\_seq6

CCGCTCTCCTTCGCACATAGGGACACGCCTCCCAGGAAGGTTTGCACACATCATATCCTCTAACAAACATATTTT

CTTTCA  
TTCTTATTTACATGACAAATATAACACATAGTCATCAAAAATTGAGGGATCGCAATCAGAGGCCTTCACTTATA  
TTTTCA  
TTTTATTCATTACAAGATCTCAACACTTGGTTCGCATTTTGGTCTCAGTGTTATTTTAATCACATTACTCCTAA  
AACATA  
TTTCATAAACATGGCCTTCAAGAACTTGCCAGACCAAGGATGACAATGACAATCATTTCACAAACGCCCTAAGC  
AGTCAT  
TTCACGACGCTCTAAGCAGTCGTTGTAAACTTGTAACAATACGGCTCTTATTAGTCTTACTCTCAACGGGCTT  
AAGCCA TCCTGCAGTAGGCTAATAAA

Horizontal gene transfer of phytochelatin synthases from bacteria to extremophilic green algae; Microbial Ecology;  
Olsson S\*, Penacho V, Puente-Sánchez F, Díaz S, Aguilera A;  
\* corresponding author, Department of Agricultural Sciences, P.O. Box 27,  
00014 University of Helsinki, Finland, sanna.olsson@helsinki.fi.

Online Resource 7 Isoforms of component compl1852\_c0 coding for the CaPCS2 gene in *Chlamydomonas acidophila*

```
>compl1852_c0_seq1_len=2482
TTTGTTCACCTTTTCGGCAGTAAATAAGCAATATTGTAGTCAAGTTGTTATCCTCGCTCTTACCTTTTGATTTC
AGGATG
CTTTCAGTCTTGCTTTGTACCAGAATTTTGAAAACATACATTCTCTTCAAGTTCTGACCTCATCCGTCCAAGG
GGAGTT
GGGTCTGCCTTGTTCCCTTTTGTCAATTAACATCACACGAAGGACAAAAGGCCTTCATTTCGTGCCAAGTATAAGG
GCGATG
CATTTTCCATGCTGGAGAACTTTGTTGCACAGGTCAACGATGCGTACTGCGGGCCAGCCTCTGCTGTAACAGTC
CTGAAT
TCTTTGGGGTGGGGCTCCCTATGCAGGACCCTGCTTCAGGGTTAGACTTGCCTGAATTTGCATATTACACACAA
GACAAC
GTATTCTCACTAGAAACAGAAAGAGTGAAGCCCAAAGCTCAGATCAGGGAAGAGGGCCTCAGCCTATCAGAGCT
CTCAAG
TTTCATCGCTGCTCACTCAGGAGTCAACGCCACGTACCGCCACAGCAGGGACATGAAAGACATGAATGAGTTTC
GACAGC
TGGCTGTGTCAGCTCTAGCACAGCCTAATGTGTATGTTATTGTCAACTTCTTCCGGCAGACATTGCAGGAACGA
GGCCAG
GGTCATCACTCTCCAATTGCAGCATACAATGAAGAAAAAGACTCTTTCCTGGTATTGGATGTCTCCAGATACAA
GTATCC
ATCTTGGTGGGTACCAGTTTCAGAGCTGTTTGCAGCAGTCCTTCCAAGAGATCCGTCAGATGTAGAAGATCCAG
GAAGGG
GCCTCGTGTTTGTCTGGGCTCAAACATCAGATGGGACACTGCAACCTTCCTTAGGAAGGGAGTTCTTCTCATCA
GGGCTG
GACGATGGCATGCCTACTTATGACTTTTGCATGATGTGCACTCAGCATTGTTGGTCTTGAGGTGCTGTCCTTTTG
GCTGGG
ATTGGGATATTGTGCTTGCTCTGTTGTTGTTTATGTCAACTGGAGTTATCAAGGTCAGCTACCAGGCACCCTC
TTGGGG
GATCGTTGGAATAAGAACTTCTGGGCTAGGGACCTGAGATGTTCAATCAAAGTGCAGGGGCATAAACAGATTG
CTACCG
AATTCCATCAAGACCAGCAGCGAAGAGCAGGCGCTGACTATGAGTTATCTGCAACATCATTTGATGACACTGAA
TTATCA
CCACTGTGCGCTGTTAGATGAACTGAAGCGGTTCTTCTGATTCTTTGCATGTCAGCTTTTAAACGTTCTTGACA
AAACTG
TTCTTGATGTCATCATGCGTGCCCTCGGTTTTTACATGAGCACAAGCTCTTCGAATCATCAATGTAAAGAGTAC
ATTGAA
TAGCATAGATAGGAAAACACAAAAATTGAACATATTTTTTGGATAATCAGGCTATCGAGAGCAACTGGTGGA
AAAAGT
TATACATACTATCTCTGTCTCAATACAACAACATTTTTTATGAAGAAACAGAAGCCTTGTAACATAGAAAGCAG
GAAATG
TTCATCATCTCCAATTAATTCTGAGGTCTGATGTAGCAAGCTTCACCAAAGAGGTAAGAACATGCAGTTCCAAG
GATTCT
GTTCTGACCACTCTCCTGCTCAGGTTTTGATTTTCATGGCCTCTGCCACTGGAGAAGCGACAGCTCCGTTGTTA
TTGCTG
GTAGCGTTCAGCTTGTTTGACACCTCATCAATGGAGGCGTTGAGCTGTGCAATCTTATCAAACAGATCCTGTTT
GGTCTG
TTCAGGGTCCTTCACCTCATCCTCAAGAAATCTTGGAAGTTTCAGGCGTTGATCATCGCCCAAGAGCGCTTTAG
AGATCT
GCAGAGAACAAAATAAAACCGCATGGTTCACACACCGCGGGAAAGAACGGCACATGGAATGAATGACATGGCAT
```

GTTAAT  
AAGCACAGTTAAATAAATCTCCATGTTGAGGTATCTCAATTTCAAGATTATGATGTAGGTAACTGCGTTTCAGC  
AGTTAT  
TCAAAAAATTAGTAGGCCTACATGGCTATTGCACATGTGTATGTGATGCCTTTGCAATTCATCATACCTGCGGG  
GCAAAAT  
ATGAATCCCAGAGTACCGAAAACAACACCCCCAAGGACAAATCCAGAGACGAAGCTGTCACTTCGGTCACTATT  
CTCAGC  
TCTGCAAGCGACAATAATGAGATCAACAACACAAATCCAATCATGAAGGTATGAGCATGTCTCTCACCTAGCTA  
CGATAT  
TCCTCTCTTTTCTTAACGTTCTTGTACAGGAGTATGCTGTATCGGAGAGCTTTTCGCGACAATGGAAAGATCCC  
GAACAT  
AAGTTCCTTGAGGTCAAACAAAAGATGGGAGGAGAATTCATCGGGGTACAAAACGTGCAGCTTACCTGTCTGC  
TTGTTA  
ACATTCTTGAACACTACTATGCGTATCGTTAATCAGGAGGACTCCGTGTAAGCGTTAAAGGGCAATCATGCAGC  
CACGTA AG

>comp11852\_c0\_seq2\_len=2416

TTTGTTCACCTTTTCGGCAGTAAATAAGCAATATTGTAGTCAAGTTGTTATCCTCGCTCTTACCTTTTGATTTC  
AGGATG  
CTTTCAGTCTTGCTTTGTACCAGAATTTTGAAAACATACATTCTCTTTCAGTTCTGACCTCATCCGTCCAAGG  
GGAGTT  
GGGCTGCTGCTTGTTCCTTTTGTCAATTAACATCACACGAAGGACAAAAGGCCTTCATTTCGTGCCAAGTATAAGG  
GCGATG  
CATTTTCCATGCTGGAGAACTTTGTTGCACAGGTCAACGATGCGTACTGCGGGCCAGCCTCTGCTGTAACAGTC  
CTGAAT  
TCTTTGGGGTGGGGCTCCCTATGCAGGACCCTGCTTCAGGGTTAGACTTGCCTGAATTTGCATATTACACACAA  
GACAAAC  
GTATTCTCACTAGAAACAGAAAGAGTGAAGCCCAAAGCTCAGATCAGGGAAGAGGGCCTCAGCCTATCAGAGCT  
CTCAAG  
TTTCATCGCTGCTCACTCAGGAGTCAACGCCACGTACCGCCACAGCAGGGACATGAAAGACATGAATGAGTTTC  
GACAGC  
TGGCTGTGTCAGCTCTAGCACAGCCTAATGTGTATGTTATTGTCAACTTCTTCCGGCAGACATTGCAGGAACGA  
GGCCAG  
GGTCATCACTCTCCAATTGCAGCATAACAATGAAGAAAAAGACTCTTTCCTGGTATTGGATGTCTCCAGATACAA  
GTATCC  
ATCTTGGTGGGTACCAGTTTCAGAGCTGTTTGCAGCAGTCCTTCCAAGAGATCCGTCAGATGTAGAAGATCCAG  
GAAGGG  
GCCTCGTGTTTGTCTGGGCTCAAACATCAGATGGGACACTGCAACCTTCCTTAGGAAGGGAGTTCTTCTCATCA  
GGGCTG  
GACGATGGCATGCCTACTTATGACTTTTGCATGATGTGCACTCAGCATTGTTGTTGAGGTGCTGTCCTTTTG  
GCTGGG  
ATTGGGATATTGTGCTTGCCTCTGTTGTTGTTTATGTCAACTGGAGTTATCAAGGTCAGCTACCAGGCACCCTC  
TTGGGG  
GATCGTTGGAATAAGAACTTCTGGGCTAGGGACCTGAGATGTTCAATCAAAGTGCAGGGGCATAAACAGATTG  
CTACCG  
AATTCCATCAAGACCAGCAGCGAAGAGCAGGCGCTGACTATGAGTTATCTGCAACATCATTTGATGACACTGAA  
TTATCA  
CCACTGTGCGCTGTTAGATGAACTGAAGCGGTCTTCTGATTCCTTTGCATGTCAGCTTTTAAACGTTCTTGACA  
AAACTG  
TTCCTGATGTCATCATGCGTGCCCTCGGTTTTTACATGAGCACAAGCTCTTCGAATCATCAATGTAAAGAGTAC  
ATTGAA  
TAGCATAGATAGGAAAACCTCACAAAAATTGAACATATTTTTTTGGATAATCAGGCTATCGAGAGCAACTGGTGGA  
AAAAGT  
TATACATACTATCTCTGTCTCAATACAACAACATTTTTTATGAAGAAACAGAAGCCTTGTAACATAGAAAGCAG  
GAAATG  
TTCATCATCTCCAATTAATTCTGAGGTCTGATGTAGCAAGCTTCACCAAAGAGGTAAGAACATGCAGTTCCAAG  
GATTCT



GTTCTGACCACTCTCCTGCTCAGGTTTTGATTTTCATGGCCTCTGCCACTGGAGAAGCGACAGCTCCGTTGTTA  
TTGCTG  
GTAGCGTTCAGCTTGTTTGACACCTCATCAATGGAGGCGTTGAGCTGTGCAATCTTATCAAACAGATCCTGTTT  
GGTCTG  
TTCAGGGTCTTTCACCTCATCCTCAAGAAATCTTGGAAGTTTCAGGCGTTGATCATCGCCCAAGAGCGCTTTAG  
AGATCT  
GCAGAGAACAAAATAAAACCGCATGGTTCACACACCGCGGGAAAGAACGGCACATGGAATGAATGACATGGCAT  
GTTAAT  
AAGCACAGTTAAATAAATCTCCATGTTGAGGTATCTCAATTTCAAGATTATGATGTAGGTAAC TGCGTTTCAGC  
AGTTAT  
TCAAAAAATTAGTAGGCCTACATGGCTATTGCACATGTGTATGTGATGCCTTTGCAATTCATCATACCTGCGGG  
GCAAT  
ATGAATCCCAGAGTACCGAAAACAACACCCCCAAGGACAAATCCAGAGACGAAGCTGTCACTTCGGTCACTATT  
CTCAGC  
TCTAGCTACGATATTCCTCTCTTTTCTTAACGTTCTTGTCACAGGAGTATGCTGTATCGGAGAGCTTTCGCGAC  
AATGGA  
AAGATCCCGAACATAAGTTCCTGAGGTCAAACAAAAGATGGGAGGAGAATTCATCGGGGTACAAAACGTGCAG  
CTTACC  
TGTCTGCTTGTTAACATTCTTGAACACTACTATGCGTATCGTTAATCAGGAGGACTCCGTGTAAGCGTTAAAG  
GGCAAT CATGCAGCCACGTAAG

>comp11852\_c0\_seq3\_len=2253

TTTGTTCACTTTTCGGCAGTAAATAAGCAATATTGTAGTCAAGTTGTTATCCTCGCTCTTACCTTTTGATTTC  
AGGATG  
CTTTCAGTCTTGCTTTGTACCAGAATTTTGAAAACATACATTCTCTTCAAGTTCTGACCTCATCCGTCCAAGG  
GGAGTT  
GGGTCTGCCTTGTTCCCTTTTGTCAATTAACATCACACGAAGGACAAAAGGCCTTCATTTCGTGCCAAGTATAAGG  
GCGATG  
CATTTTCCATGCTGGAGAACTTTGTTGCACAGGTCAACGATGCGTACTGCGGGCCAGCCTCTGCTGTAACAGTC  
CTGAAT  
TCTTTGGGGTGGGGCTCCCTATGCAGGACCCTGCTTCAGGGTTAGACTTGCCTGAATTTGCATATTACACACAA  
GACAAAC  
GTATTCTCACTAGAAACAGAAAGAGTGAAGCCCAAAGCTCAGATCAGGGAAGAGGGCCTCAGCCTATCAGAGCT  
CTCAAG  
TTTCATCGCTGCTCACTCAGGAGTCAACGCCACGTACCGCCACAGCAGGGACATGAAAGACATGAATGAGTTTC  
GACAGC  
TGGCTGTGTCAGCTCTAGCACAGCCTAATGTGTATGTTATTGTCAACTTCTTCCGGCAGACATTGCAGGAACGA  
GGCCAG  
GGTCATCACTCTCCAATTGCAGCATAACAATGAAGAAAAAGACTCTTTCCTGGTATTGGATGTCTCCAGATACAA  
GTATCC  
ATCTTGGTGGGTACCAGTTTCAGAGCTGTTTGCAGCAGTCCTTCCAAGAGATCCGTCAGATGTAGAAGATCCAG  
GAAGGG  
GCCTCGTGTTTGTCTGGGCTCAAACATCAGATGGGACACTGCAACCTTCCTTAGGAAGGGAGTTCTTCTCATCA  
GGGCTG  
GACGATGGCATGCCTACTTATGACTTTTGCATGATGTGCACTCAGCATTTGGTCTTGAGGTGCTGTCCTTTTG  
GCTGGG  
ATTGGGATATTGTGCTTGCCTCTGTTGTTGTTTATGTCAACTGGAGTTATCAAGGTCAGCTACCAGGCACCCTC  
TTGGGG  
GATCGTTGGAATAAGAACTTCTGGGCTAGGGACCTGAGATGTTCAATCAAAGTGCAGGGGCATAAACAGATTG  
CTACCG  
AATTCATCAAGACCAGCAGCGAAGAGCAGGCGCTGACTATGAGTTATCTGCAACATCATTTGATGACACTGAA  
TTATCA  
CCACTGTGCGCTGTTAGATGAACTGAAGCGGTTCTTCTGATTCCTTTGCATGTCAGCTTTTAACGTTCTTGACA  
AAACTG  
TTCCTGATGTCATCATGCGTGCCCTCGGTTTTTACATGAGCACAAGCTCTTCGAATCATCAATGTAAAGAGTAC  
ATTGAA  
TAGCATAGATAGGAAAAC TCACAAAAATTGAACATATTTTTTTGGATAATCAGGCTATCGAGAGCAACTGGTGGA

AAAAGT  
TATACATACTATCTCTGTCTCAATACAACAACATTTTTATGAAGAAACAGAAGCCTTGTAACATAGAAAGCAG  
GAAATG  
TTCATCATCTCCAATTAATTCTGAGGTCTGATGTAGCAAGCTTCACCAAAGAGGTAAGAACATGCAGTTCCAAG  
GATTCT  
GTTCTGACCACTCTCCTGCTCAGGTTTTGATTTTCATGGCCTCTGCCACTGGAGAAGCGACAGCTCCGTTGTTA  
TTGCTG  
GTAGCGTTCAGCTTGTTTGACACCTCATCAATGGAGGCGTTGAGCTGTGCAATCTTATCAAACAGATCCTGTTT  
GGTCTG  
TTCAGGGTCCTTCACCTCATCCTCAAGAAATCTTGGAAGTTTCAGGCGTTGATCATCGCCCAAGAGCGCTTTAG  
AGATCT  
GCGGGGCAAATATGAATCCCAGAGTACCGAAAACAACACCCCCAAGGACAAATCCAGAGACGAAGCTGTCACTT  
CGGTCA  
CTATTCTCAGCTCTGCAAGCGACAATAATGAGATCAACAACACAAATCCAATCATGAAGGTATGAGCATGTCTC  
TCACCT  
AGCTACGATATTCTCTCTTTTCTTAACGTTCTTGTACAGGAGTATGCTGTATCGGAGAGCTTTCGCGACAAT  
GGAAAG  
ATCCCGAACATAAGTTCCCTGAGGTCAAACAAAAGATGGGAGGAGAATTCATCGGGGTACAAAACGTGCAGCTT  
ACCTGT  
CCTGCTTGTTAACATTCTTGAACACTACTATGCGTATCGTTAATCAGGAGGACTCCGTGTAAGCGTTAAAGGGC  
AATCAT GCAGCCACGTAAG

>comp11852\_c0\_seq4\_len=2187

TTTGTTCACTTTTCGGCAGTAAATAAGCAATATTGTAGTCAAGTTGTTATCCTCGCTCTTACCTTTTGATTTC  
AGGATG  
CTTTCAGTCTTGCTTTGTACCAGAATTTTGAAAACATACATTCTCTTTCAAGTTCTGACCTCATCCGTCCAAGG  
GGAGTT  
GGGTCTGCCTTGTTCCCTTTTGTCAATTAACATCACACGAAGGACAAAAGGCCTTCATTTCGTGCCAAGTATAAGG  
GCGATG  
CATTTTCCATGCTGGAGAACTTTGTTGCACAGGTCAACGATGCGTACTGCGGGCCAGCCTCTGCTGTAACAGTC  
CTGAAT  
TCTTTGGGGTGGGGCTCCCTATGCAGGACCCTGCTTCAGGGTTAGACTTGCCTGAATTTGCATATTACACACAA  
GACAAAC  
GTATTCTCACTAGAAACAGAAAGAGTGAAGCCCAAAGCTCAGATCAGGGAAGAGGGCCTCAGCCTATCAGAGCT  
CTCAAG  
TTTCATCGCTGCTCACTCAGGAGTCAACGCCACGTACCGCCACAGCAGGGACATGAAAGACATGAATGAGTTTC  
GACAGC  
TGGCTGTGTCAGCTCTAGCACAGCCTAATGTGTATGTTATTGTCAACTTCTTCCGGCAGACATTGCAGGAACGA  
GGCCAG  
GGTCATCACTCTCCAATTGCAGCATACAATGAAGAAAAAGACTCTTTCCTGGTATTGGATGTCTCCAGATACAA  
GTATCC  
ATCTTGTTGGGTACCAGTTTCAGAGCTGTTTGCAGCAGTCCTTCCAAGAGATCCGTCAGATGTAGAAGATCCAG  
GAAGGG  
GCCTCGTGTTTGTCTGGGCTCAAACATCAGATGGGACACTGCAACCTTCCTTAGGAAGGGAGTTCTTCTCATCA  
GGGCTG  
GACGATGGCATGCCTACTTATGACTTTTGCGATGATGTGCACTCAGCATTGTTGAGGTGCTGTCCTTTTG  
GCTGGG  
ATTGGGATATTGTGCTTGCCTCTGTTGTTGTTTATGTCAACTGGAGTTATCAAGGTCAGCTACCAGGCACCCTC  
TTGGGG  
GATCGTTGGAATAAGAACTTCTGGGCTAGGGACCTGAGATGTTCAATCAAAGTGCAGGGGCATAAACAGATTG  
CTACCG  
AATTCCATCAAGACCAGCAGCGAAGAGCAGGCGCTGACTATGAGTTATCTGCAACATCATTTGATGACACTGAA  
TTATCA  
CCACTGTGCGCTGTTAGATGAACTGAAGCGGTTCTTCTGATTCCTTTGCATGTCAGCTTTTAACGTTCTTGACA  
AAACTG  
TTCCTGATGTCATCATGCGTGCCCTCGGTTTTTACATGAGCACAAGCTCTTCGAATCATCAATGTAAAGAGTAC  
ATTGAA

TAGCATAGATAGGAAACTCACAAAAATTGAACATATTTTTTTGGATAATCAGGCTATCGAGAGCAACTGGTGGA  
AAAAGT  
TATACATACTATCTCTGTCTCAATACAACAACATTTTTATGAAGAAACAGAAGCCTTGTAACATAGAAAGCAG  
GAAATG  
TTCATCATCTCCAATTAATTCTGAGGTCTGATGTAGCAAGCTTCACCAAAGAGGTAAGAACATGCAGTTCCAAG  
GATTCT  
GTTCTGACCACTCTCCTGCTCAGGTTTTGATTTTCATGGCCTCTGCCACTGGAGAAGCGACAGCTCCGTTGTTA  
TTGCTG  
GTAGCGTTCAGCTTGTTTGACACCTCATCAATGGAGGCGTTGAGCTGTGCAATCTTATCAAACAGATCCTGTTT  
GGTCTG  
TTCAGGGTCCTTCACCTCATCCTCAAGAAATCTTGGAAGTTTCAGGCGTTGATCATCGCCCAAGAGCGCTTTAG  
AGATCT  
GCGGGGCAAATATGAATCCCAGAGTACCGAAAACAACACCCCCAAGGACAAATCCAGAGACGAAGCTGTCACTT  
CGGTCA  
CTATTCTCAGCTCTAGCTACGATATTCTCTCTTTTTCTTAACGTTCTTGTCACAGGAGTATGCTGTATCGGAGA  
GCTTTC  
GCGACAATGGAAAGATCCCGAACATAAGTTCCCTGAGGTCAAACAAAAGATGGGAGGAGAATTCATCGGGGTAC  
AAAACG  
TGCAGCTTACCTGTCCTGCTTGTTAACATTCTTGAACACTACTATGCGTATCGTTAATCAGGAGGACTCCGTGT  
AAGCGT TAAAGGGCAATCATGCAGCCACGTAAG

>comp11852\_c0\_seq5\_len=2129

TTTGTTCACTTTTCGGCAGTAAATAAGCAATATTGTAGTCAAGTTGTTATCCTCGCTCTTACCTTTTGATTTC  
AGGATG  
CTTTCAGTCTTGCTTTGTACCAGAATTTTGAAAACATACATTCTCTTCAAGTTCTGACCTCATCCGTCCAAGG  
GGAGTT  
GGGTCTGCCTTGTTCCCTTTTGTCAATTAACATCACACGAAGGACAAAAGGCCTTCATTTCGTGCCAAGTATAAGG  
GCGATG  
CATTTTCCATGCTGGAGAACTTTGTTGCACAGGTCAACGATGCGTACTGCGGGCCAGCCTCTGCTGTAACAGTC  
CTGAAT  
TCTTTGGGGTGGGGCTCCCTATGCAGGACCCTGCTTCAGGGTTAGACTTGCCTGAATTTGCATATTACACACAA  
GACAAAC  
GTATTCTCACTAGAAACAGAAAGAGTGAAGCCCAAAGCTCAGATCAGGGAAGAGGGCCTCAGCCTATCAGAGCT  
CTCAAG  
TTTCATCGCTGCTCACTCAGGAGTCAACGCCACGTACCGCCACAGCAGGGACATGAAAGACATGAATGAGTTTC  
GACAGC  
TGGCTGTGTCAGCTCTAGCACAGCCTAATGTGTATGTTATTGTCAACTTCTTCCGGCAGACATTGCAGGAACGA  
GGCCAG  
GGTCATCACTCTCCAATTGCAGCATAACAATGAAGAAAAAGACTCTTTCCTGGTATTGGATGTCTCCAGATACAA  
GTATCC  
ATCTTGGTGGGTACCAGTTTCAGAGCTGTTTGCAGCAGTCCTTCCAAGAGATCCGTCAGATGTAGAAGATCCAG  
GAAGGG  
GCCTCGTGTTTGTCTGGGCTCAAACATCAGATGGGACACTGCAACCTTCCTTAGGAAGGGAGTTCTTCTCATCA  
GGGCTG  
GACGATGGCATGCCTACTTATGACTTTTGCATGATGTGCACTCAGCATTTGGTCTTGAGGTGCTGTCCTTTTG  
GCTGGG  
ATTGGGATATTGTGCTTGCCTCTGTTGTTGTTTATGTCAACTGGAGTTATCAAGGTCAGCTACCAGGCACCCTC  
TTGGGG  
GATCGTTGGAATAAGAACTTCTGGGCTAGGGACCTGAGATGTTCAATCAAAGTGCAGGGGCATAAACAGATTG  
CTACCG  
AATTCATCAAGACCAGCAGCGAAGAGCAGGCGCTGACTATGAGTTATCTGCAACATCATTTGATGACACTGAA  
TTATCA  
CCACTGTGCGCTGTTAGATGAACTGAAGCGGTTCTTCTGATTCCTTTGCATGTCAGCTTTTAACGTTCTTGACA  
AAACTG  
TTCCTGATGTCATCATGCGTGCCCTCGGTTTTTACATGAGCACAAGCTCTTCGAATCATCAATGTAAAGAGTAC  
ATTGAA  
TAGCATAGATAGGAAACTCACAAAAATTGAACATATTTTTTTGGATAATCAGGCTATCGAGAGCAACTGGTGGA

AAAAGT  
TATACATACTATCTCTGTCTCAATACAACAACATTTTTATGAAGAAACAGAAGCCTTGTAACATAGAAAGCAG  
GAAATG  
TTCATCATCTCCAATTAATTCTGAGGTCTGATGTAGCAAGCTTCACCAAAGAGGTAAGAACATGCAGTTCCAAG  
GATTCT  
GTTCTGACCACTCTCCTGCTCAGGTTTTGATTTTCATGGCCTCTGCCACTGGAGAAGCGACAGCTCCGTTGTTA  
TTGCTG  
GTAGCGTTCAGCTTGTTTGACACCTCATCAATGGAGGCGTTGAGCTGTGCAATCTTATCAAACAGATCCTGTTT  
GGTCTG  
TTCAGGGTCCTTCACCTCATCCTCAAGAAATCTTGGAAGTTTCAGGCGTTGATCATCGCCCAAGAGCGCTTTAG  
AGATCT  
GCGGGGCAAATATGAATCCCAGAGTACCGAAAACAACACCCCCAAGGACAAATCCAGAGACGAAGCTGTCACTT  
CGGTCA  
CTATTCTCAGCTCTAGCTACGATATTCTCTCTTTTTCTTAACGTTCTTGTCACAGGAGTATGCTGTATCGGAGA  
GCTTTC  
GCGACAATGGAAAGATCCCGAACATAAGTTCCTGTCTGCTTGTTAACATTCTTGAACACTACTATGCGTATC  
GTTAAT CAGGAGGACTCCGTGTAAGCGTTAAAGGGCAATCATGCAGCCACGTAAG

>comp11852\_c0\_seq6\_len=2424

TTTGTTCACTTTTTCGGCAGTAAATAAGCAATATTGTAGTCAAGTTGTTATCCTCGCTCTTACCTTTTGATTTC  
AGGATG  
CTTTCAGTCTTGCTTTGTACCAGAATTTTGAAAACATACATTCTCTTCAAGTTCTGACCTCATCCGTCCAAGG  
GGAGTT  
GGGTCTGCCTTGTTCCCTTTTGTCAATTAACATCACACGAAGGACAAAAGGCCTTCATTTCGTGCCAAGTATAAGG  
GCGATG  
CATTTTCCATGCTGGAGAACTTTGTTGCACAGGTCAACGATGCGTACTGCGGGCCAGCCTCTGCTGTAACAGTC  
CTGAAT  
TCTTTGGGGTGGGGCTCCCTATGCAGGACCCTGCTTCAGGGTTAGACTTGCCTGAATTTGCATATTACACACAA  
GACAAC  
GTATTCTCACTAGAAACAGAAAGAGTGAAGCCCAAAGCTCAGATCAGGGAAGAGGGCCTCAGCCTATCAGAGCT  
CTCAAG  
TTTCATCGCTGCTCACTCAGGAGTCAACGCCACGTACCGCCACAGCAGGGACATGAAAGACATGAATGAGTTTC  
GACAGC  
TGGCTGTGTCTAGCTCTAGCACAGCCTAATGTGTATGTTATTGTCAACTTCTTCCGGCAGACATTGCAGGAACGA  
GGCCAG  
GGTCATCACTCTCCAATTGCAGCATACAATGAAGAAAAAGACTCTTTCCTGGTATTGGATGTCTCCAGATACAA  
GTATCC  
ATCTTGGTGGGTACCAGTTTCAGAGCTGTTTGCAGCAGTCCTTCCAAGAGATCCGTCAGATGTAGAAGATCCAG  
GAAGGG  
GCCTCGTGTTTGTCTGGGCTCAAACATCAGATGGGACACTGCAACCTTCCTTAGGAAGGGAGTTCTTCTCATCA  
GGGCTG  
GACGATGGCATGCCTACTTATGACTTTTGCATGATGTGCACTCAGCATTGTTGAGGTGCTGTCCTTTTG  
GCTGGG  
ATTGGGATATTGTGCTTGCCTCTGTTGTTGTTTATGTCAACTGGAGTTATCAAGGTCAGCTACCAGGCACCCTC  
TTGGGG  
GATCGTTGGAATAAGAACTTCTGGGCTAGGGACCTGAGATGTTCAATCAAAGTGCAGGGGCATAAACAGATTG  
CTACCG  
AATTCCATCAAGACCAGCAGCGAAGAGCAGGCGCTGACTATGAGTTATCTGCAACATCATTGATGACACTGAA  
TTATCA  
CCACTGTGCGCTGTTAGATGAACTGAAGCGGTTCTTCTGATTCTTTGCATGTCAGCTTTTAACGTTCTTGACA  
AAACTG  
TTCCTGATGTCATCATGCGTGCCCTCGGTTTTTACATGAGCACAAGCTCTTCGAATCATCAATGTAAAGAGTAC  
ATTGAA  
TAGCATAGATAGGAAACTCACAAAATTTGAACATATTTTTTGGATAATCAGGCTATCGAGAGCAACTGGTGGA  
AAAAGT  
TATACATACTATCTCTGTCTCAATACAACAACATTTTTATGAAGAAACAGAAGCCTTGTAACATAGAAAGCAG  
GAAATG

TTCATCATCTCCAATTAATTCTGAGGTCTGATGTAGCAAGCTTCACCAAAGAGGTAAGAACATGCAGTTCCAAG  
GATTCT  
GTTCTGACCACTCTCCTGCTCAGGTTTTGATTTTCATGGCCTCTGCCACTGGAGAAGCGACAGCTCCGTTGTTA  
TTGCTG  
GTAGCGTTCAGCTTGTTTGACACCTCATCAATGGAGGCGTTGAGCTGTGCAATCTTATCAAACAGATCCTGTTT  
GGTCTG  
TTCAGGGTCCTTCACCTCATCCTCAAGAAATCTTGGAAGTTTCAGGCGTTGATCATCGCCCAAGAGCGCTTTAG  
AGATCT  
GCAGAGAACAAAATAAAACCGCATGGTTCACACACCGCGGGAAAGAACGGCACATGGAATGAATGACATGGCAT  
GTTAAT  
AAGCACAGTTAAATAAATCTCCATGTTGAGGTATCTCAATTTCAAGATTATGATGTAGGTAAGTGCCTTTCAGC  
AGTTAT  
TCAAAAAATTAGTAGGCCTACATGGCTATTGCACATGTGTATGTGATGCCTTTGCAATTCATCATACCTGCGGG  
GCAAAT  
ATGAATCCCAGAGTACCGAAAACAACACCCCCAAGGACAAATCCAGAGACGAAGCTGTCACTTCGGTCACTATT  
CTCAGC  
TCTGCAAGCGACAATAATGAGATCAACAACACAAATCCAATCATGAAGGTATGAGCATGTCTCTCACCTAGCTA  
CGATAT  
TCCCTCTCTTTTCTTAACGTTCTTGTCACAGGAGTATGCTGTATCGGAGAGCTTTTCGCGACAATGGAAAGATCCC  
GAACAT  
AAGTTCCCTGTCTGCTTGTTAACATTCTTGAACACTACTATGCGTATCGTTAATCAGGAGGACTCCGTGTAAG  
CGTTAA AGGGCAATCATGCAGCCACGTAAG

>comp11852\_c0\_seq7\_len=2195

TTTGTTCACCTTTTCGGCAGTAAATAAGCAATATTGTAGTCAAGTTGTTATCCTCGCTCTTACCTTTTGATTTC  
AGGATG  
CTTTCAGTCTTGCTTTGTACCAGAATTTTGAAAACATACATTCTCTTTCAAGTTCTGACCTCATCCGTCCAAGG  
GGAGTT  
GGGTCTGCCTTGTTCCCTTTTGTCAATTAACATCACACGAAGGACAAAAGGCCTTCATTTCGTGCCAAGTATAAGG  
GCGATG  
CATTTTCCATGCTGGAGAACTTTGTTGCACAGGTCAACGATGCGTACTGCGGGCCAGCCTCTGCTGTAACAGTC  
CTGAAT  
TCTTTGGGGTGGGGCTCCCTATGCAGGACCCTGCTTCAGGGTTAGACTTGCCTGAATTTGCATATTACACACAA  
GACAAAC  
GTATTCTCACTAGAAACAGAAAGAGTGAAGCCCAAAGCTCAGATCAGGGAAGAGGGCCTCAGCCTATCAGAGCT  
CTCAAG  
TTTCATCGCTGCTCACTCAGGAGTCAACGCCACGTACCGCCACAGCAGGGACATGAAAGACATGAATGAGTTTC  
GACAGC  
TGGCTGTGTCAGCTCTAGCACAGCCTAATGTGTATGTTATTGTCAACTTCTTCCGGCAGACATTGCAGGAACGA  
GGCCAG  
GGTCATCACTCTCCAATTGCAGCATACAATGAAGAAAAAGACTCTTTCCTGGTATTGGATGTCTCCAGATACAA  
GTATCC  
ATCTTGGTGGGTACCAGTTTCAGAGCTGTTTGCAGCAGTCCTTCCAAGAGATCCGTGAGATGTAGAAGATCCAG  
GAAGGG  
GCCTCGTGTTTGTCTGGGCTCAAACATCAGATGGGACACTGCAACCTTCCTTAGGAAGGGAGTTCTTCTCATCA  
GGGCTG  
GACGATGGCATGCCTACTTATGACTTTTGCATGATGTGCACTCAGCATTTGGTCTTGAGGTGCTGTCCTTTTG  
GCTGGG  
ATTGGGATATTGTGCTTGCCTCTGTTGTTGTTTATGTCAACTGGAGTTATCAAGGTCAGCTACCAGGCACCCTC  
TTGGGG  
GATCGTTGGAATAAGAACTTCTGGGCTAGGGACCTGAGATGTTCAATCAAAGTGCAGGGGCATAAACAGATTG  
CTACCG  
AATTCATCAAGACCAGCAGCGAAGAGCAGGCGCTGACTATGAGTTATCTGCAACATCATTTGATGACACTGAA  
TTATCA  
CCACTGTGCGCTGTTAGATGAACTGAAGCGGTTCTTCTGATTCCTTTGCATGTCAGCTTTTAACGTTCTTGACA  
AAACTG  
TTCTGATGTCATCATGCGTGCCCTCGGTTTTTACATGAGCACAAGCTCTTCGAATCATCAATGTAAAGAGTAC

ATTGAA  
TAGCATAGATAGGAAACTCACAAAAATTGAACATATTTTTTTGGATAATCAGGCTATCGAGAGCAACTGGTGGA  
AAAAGT  
TATACATACTATCTCTGTCTCAATACAACAACATTTTTTATGAAGAAACAGAAGCCTTGTAACATAGAAAGCAG  
GAAATG  
TTCATCATCTCCAATTAATTCTGAGGTCTGATGTAGCAAGCTTCACCAAAGAGGTAAGAACATGCAGTTCCAAG  
GATTCT  
GTTCTGACCACTCTCCTGCTCAGGTTTTGATTTTCATGGCCTCTGCCACTGGAGAAGCGACAGCTCCGTTGTTA  
TTGCTG  
GTAGCGTTCAGCTTGTTTGACACCTCATCAATGGAGGCGTTGAGCTGTGCAATCTTATCAAACAGATCCTGTTT  
GGTCTG  
TTCAGGGTCCTTCACCTCATCCTCAAGAAATCTTGGAAGTTTCAGGCGTTGATCATCGCCCAAGAGCGCTTTAG  
AGATCT  
GCGGGGCAAATATGAATCCCAGAGTACCGAAAAACAACACCCCCAAGGACAAATCCAGAGACGAAGCTGTCACTT  
CGGTCA  
CTATTCTCAGCTCTGCAAGCGACAATAATGAGATCAACAACACAAATCCAATCATGAAGGTATGAGCATGTCTC  
TCACCT  
AGCTACGATATTCCTCTCTTTTCTTAACGTTCTTGTCACAGGAGTATGCTGTATCGGAGAGCTTTCGCGACAAT  
GGAAAG  
ATCCCGAACATAAGTTCCCTGTCCTGCTTGTTAACATTCTTGAACACTACTATGCGTATCGTTAATCAGGAGGA  
CTCCGT GTAAGCGTTAAAGGGCAATCATGCAGCCACGTAAG

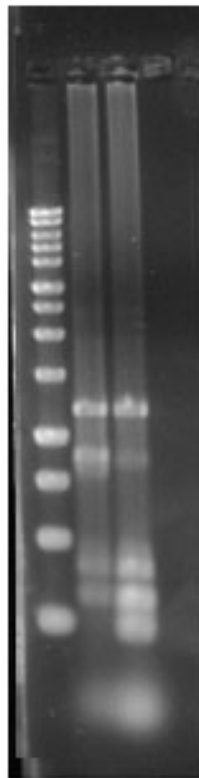
>comp11852\_c0\_seq8\_len=2358

TTTGTTCACTTTTCGGCAGTAAATAAGCAATATTGTAGTCAAGTTGTTATCCTCGCTCTTACCTTTTGATTTC  
AGGATG  
CTTTCAGTCTTGCTTTGTACCAGAATTTTGAAAACATACATTCTCTTTCAAGTTCTGACCTCATCCGTCCAAGG  
GGAGTT  
GGGTCTGCCTTGTTCCCTTTTGTCAATTAACATCACACGAAGGACAAAAGGCCTTCATTTCGTGCCAAGTATAAGG  
GCGATG  
CATTTTCCATGCTGGAGAACTTTGTTGCACAGGTCAACGATGCGTACTGCGGGCCAGCCTCTGCTGTAACAGTC  
CTGAAT  
TCTTTGGGGTGGGGCTCCCTATGCAGGACCCTGCTTCAGGGTTAGACTTGCCTGAATTTGCATATTACACACAA  
GACAAAC  
GTATTCTCACTAGAAACAGAAAGAGTGAAGCCCAAAGCTCAGATCAGGGAAGAGGGCCTCAGCCTATCAGAGCT  
CTCAAG  
TTTCATCGCTGCTCACTCAGGAGTCAACGCCACGTACCGCCACAGCAGGGACATGAAAGACATGAATGAGTTTC  
GACAGC  
TGGCTGTGTCAGCTCTAGCACAGCCTAATGTGTATGTTATTGTCAACTTCTTCCGGCAGACATTGCAGGAACGA  
GGCCAG  
GGTCATCACTCTCCAATTGCAGCATACAATGAAGAAAAAGACTCTTTCCTGGTATTGGATGTCTCCAGATACAA  
GTATCC  
ATCTTGGTGGGTACCAGTTTCAGAGCTGTTTGCAGCAGTCCTTCCAAGAGATCCGTCAGATGTAGAAGATCCAG  
GAAGGG  
GCCTCGTGTTTGTCTGGGCTCAAACATCAGATGGGACACTGCAACCTTCCTTAGGAAGGGAGTTCTTCTCATCA  
GGGCTG  
GACGATGGCATGCCTACTTATGACTTTTGCGATGATGTGCACTCAGCATTTGGTCTTGAGGTGCTGTCCTTTTG  
GCTGGG  
ATTGGGATATTGTGCTTGCCTCTGTTGTTGTTTATGTCAACTGGAGTTATCAAGGTCAGCTACCAGGCACCCTC  
TTGGGG  
GATCGTTGGAATAAGAACTTCTGGGCTAGGGACCTGAGATGTTCAATCAAAGTGCAGGGGCATAAACAGATTG  
CTACCG  
AATTCCATCAAGACCAGCAGCGAAGAGCAGGCGCTGACTATGAGTTATCTGCAACATCATTTGATGACACTGAA  
TTATCA  
CCACTGTGCGCTGTTAGATGAACTGAAGCGGTTCTTCTGATTCCTTTGCATGTCAGCTTTTAACGTTCTTGACA  
AAACTG  
TTCCTGATGTCATCATGCGTGCCCTCGGTTTTTACATGAGCACAAGCTCTTCGAATCATCAATGTAAAGAGTAC  
ATTGAA

TAGCATAGATAGGAAACTCACAAAAATTGAACATATTTTTTTGGATAATCAGGCTATCGAGAGCAACTGGTGGA  
AAAAGT  
TATACATACTATCTCTGTCTCAATACAACAACATTTTTTATGAAGAAACAGAAGCCTTGTAACATAGAAAGCAG  
GAAATG  
TTCATCATCTCCAATTAATTCTGAGGTCTGATGTAGCAAGCTTCACCAAAGAGGTAAGAACATGCAGTTCCAAG  
GATTCT  
GTTCTGACCACTCTCCTGCTCAGGTTTTGATTTTCATGGCCTCTGCCACTGGAGAAGCGACAGCTCCGTTGTTA  
TTGCTG  
GTAGCGTTCAGCTTGTTTGACACCTCATCAATGGAGGCGTTGAGCTGTGCAATCTTATCAAACAGATCCTGTTT  
GGTCTG  
TTCAGGGTCCTTCACCTCATCCTCAAGAAATCTTGGAAGTTTCAGGCGTTGATCATCGCCCAAGAGCGCTTTAG  
AGATCT  
GCAGAGAACAAAATAAAACCGCATGGTTCACACACCGCGGGAAAGAACGGCACATGGAATGAATGACATGGCAT  
GTTAAT  
AAGCACAGTTAAATAAATCTCCATGTTGAGGTATCTCAATTTCAAGATTATGATGTAGGTAACTGCGTTTCAGC  
AGTTAT  
TCAAAAAATTAGTAGGCCTACATGGCTATTGCACATGTGTATGTGATGCCTTTGCAATTCATCATACCTGCGGG  
GCAAAT  
ATGAATCCCAGAGTACCGAAAACAACACCCCCAAGGACAAATCCAGAGACGAAGCTGTCACTTCGGTCACTATT  
CTCAGC  
TCTAGCTACGATATTCCTCTCTTTTCTTAACGTTCTTGTCACAGGAGTATGCTGTATCGGAGAGCTTTCGCGAC  
AATGGA  
AAGATCCCGAACATAAGTTCCCTGTCCTGCTTGTTAACATTCTTGAACACTACTATGCGTATCGTTAATCAGGA  
GGACTC CGTGTAAGCGTTAAAGGGCAATCATGCAGCCACGTAAG

Horizontal gene transfer of phytochelatin synthases from bacteria to extremophilic green algae; Microbial Ecology; Olsson S\*, Penacho V, Puente-Sánchez F, Díaz S, Aguilera A; \* corresponding author, Department of Agricultural Sciences, P.O. Box 27, 00014 University of Helsinki, Finland, sanna.olsson@helsinki.fi

**Online Resource 8** PCR products amplified in *Dunaliella acidophila* strain RT5 using gradient annealing run on 1% agarose gel (1X TBE). The gel electrophoresis shows amplification of products sized 1198 bp and 542 bp and several non-specific amplification products. The intensity of the non-specific bands depends on the used annealing temperature and PCR program. Wells 1: Ladder 1kb (Promega), marker sizes from 250 kb to 10000 kb, well2: annealing 52-55°C. Well 3: annealing 55-60°C.





Horizontal gene transfer of phytochelatase synthases from bacteria to extremophilic green algae; Microbial Ecology; Olsson S\*, Penacho V, Puente-Sánchez F, Díaz S, Aguilera A; \* corresponding author, Department of Agricultural Sciences, P.O. Box 27, 00014 University of Helsinki, Finland, sanna.olsson@helsinki.fi

**Online Resource 9** Phylogram based on the novel phytochelatase synthase genes and the most similar genes obtained with a BLASTx search against NCBI's databases. The tree represents the majority consensus of trees sampled after stationarity in the Bayesian analysis. PP values from the Bayesian inference > 0.50 are indicated above each branch. The amino acid alignment was 383 aa long.

